(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2017-41260
(P2017-41260A)
(43)公開日　平成29年2月23日(2017.2.23)

(51) Int.Cl.			F I			テーマコード（参考）
G06F	3/01	(2006.01)	G06F	3/01	510	3C707
B25J	5/00	(2006.01)	B25J	5/00	F	5E555

審査請求 有　請求項の数 21　ＯＬ　（全 42 頁）

(21)出願番号　特願2016-186918(P2016-186918)
(22)出願日　平成28年9月26日(2016.9.26)
(62)分割の表示　特願2013-520054(P2013-520054)の分割
原出願日　平成23年7月11日(2011.7.11)
(31)優先権主張番号　1056047
(32)優先日　平成22年7月23日(2010.7.23)
(33)優先権主張国　フランス(FR)

（特許庁注：以下のものは登録商標）
１．ブルートゥース

(71)出願人　311003754
ソフトバンク・ロボティクス・ヨーロッパ
ＳＯＦＴＢＡＮＫ　ＲＯＢＯＴＩＣＳ　ＥＵＲＯＰＥ
フランス国、７５０１５・パリ、リュ・デュ・コロネル・ピエール・アビア・４３
(74)代理人　100071054
弁理士　木村　高久
(72)発明者　メゾニエ、ブルーノ
フランス、エフ－７５００５　パリ、リュジョフロワ　サン－ティレール、５７
(72)発明者　モンソー、ジェローム
フランス、エフ－７５０１３　パリ、リュブッサニョー、６

最終頁に続く

(54)【発明の名称】自然な対話インターフェースを備えたヒューマノイドロボット、同ロボットを制御する方法、および対応プログラム

(57)【要約】　　（修正有）
【課題】音声および可視認識処理の能力に制限されない、対話者との自然な対話のためのインターフェースを備えたヒューマノイドロボットを提供する。
【解決手段】ロボットは、音声信号や画像信号を受信する複数のタイプの受信チャネル５２１、５２２、５２３を備える。制御モジュール５１０は、受信チャネル５２１、５２２、５２３からメッセージを受信し、受信したメッセージの理解の信頼水準を評価する。制御モジュール５３０は、信頼水準の評価結果に依存するメッセージを生成し、出力チャネル５３１、５３２、５３３出力する。
【効果】ロボットの対話者と自然な対話を行い、対話の品質を改良する。
【選択図】図５

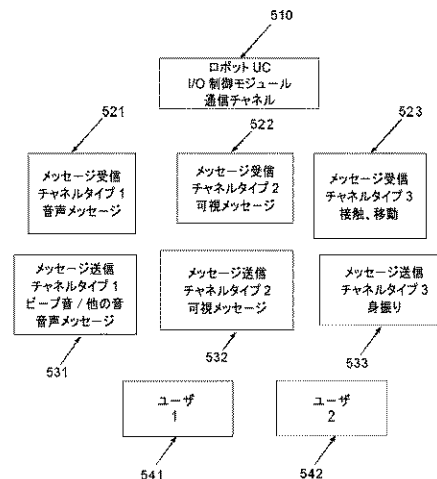



図５

【特許請求の範囲】
【請求項１】
異なる様式に応じた少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つのチャネルと、前記チャネルの入力／出力のための制御モジュールと、を含むヒューマノイドロボットであって、
前記制御モジュールは、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成される、ヒューマノイドロボットであり、
第１のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される第１のタイプのメッセージを受信する受信チャネルであり、第２のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される前記第１のタイプとは異なる第２のタイプのメッセージを送信する送信チャネルであり、
前記制御モジュールは、前記第１のタイプのメッセージを受信する受信チャネル上で受信された第１のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第２のタイプのメッセージを送信する送信チャネル上にその内容が前記信頼水準に依存する少なくとも１つの第２のメッセージを生成することができる、ヒューマノイドロボット。
【請求項２】
前記チャネルは、可聴、可視、触知可能、身振りの、位置的、または符号メッセージを送信および／または受信するための通信チャネル群から選択される、請求項１に記載のヒューマノイドロボット。
【請求項３】
第１の通信チャネルは音声送信チャネルであり、第２の通信チャネルは前記少なくとも１人の対話者による前記ロボットの少なくとも一部の身振りおよび／または位置を受信するためのチャネルであり、
前記身振りおよび／または位置は前記対話者により前記ロボットに伝えられる入力を表し、
前記入力の仕様は、前記第１のチャネル上で送信された前記メッセージにより、前記ロボットにより前記対話者に対し規定される、請求項２に記載のヒューマノイドロボット。
【請求項４】
前記対話者が前記第２のチャネル内への前記入力を検証する第３の触知能通信チャネルをさらに含む、請求項３に記載のヒューマノイドロボット。
【請求項５】
前記第１のチャネルは、それぞれが期待認識率と関連付けられた表現のリストにより、受信されたメッセージを音声認識するためのフィルタを含み、
前記第２のメッセージの前記内容は、前記第１のチャネル上の前記第１のメッセージの繰り返しの要求と、前記フィルタの前記表現のサブセットの前記第１のチャネル上で前記対話者により送信される第３のメッセージによる確認の要求と、少なくとも１つの第３のチャネル上の少なくとも１つの他のメッセージの前記対話者による送信の要求と、からなる要求の群から発見的方法により選択される、請求項２に記載のヒューマノイドロボット。
【請求項６】
前記第１と第２のチャネル上の前記メッセージの半二重モードでの順番を保証するように、前記第１のチャネル上での聴取開始信号を前記第２のチャネル上で送信するようにさらに構成される、請求項５に記載のヒューマノイドロボット。
【請求項７】
前記発見的方法は前記期待認識率に基づき決定される閾値に関する実認識率の位置の関

数である、請求項５に記載のヒューマノイドロボット。
【請求項８】
前記第３のチャネルは触覚的受信のためのまたは前記ロボットの一部から身振りを受信するためのチャネルである、請求項５に記載のヒューマノイドロボット。
【請求項９】
電子メールシステムとインターフェースするためのモジュールであって、前記第１と第２のチャネルを使用することにより、前記電子メールシステム上のアカウントのホルダが前記ロボットを、前記第２のチャネル上の電子メールを受信し／読み込み、前記第１のチャネル上で書き込み／送信し、対話により前記アカウントを管理するためのエージェントとして使用できるようにする、モジュール、をさらに含む請求項２に記載のヒューマノイドロボット。
【請求項１０】
前記第３のチャネルは前記第１のチャネルの前記フィルタの表現の前記リストに対応する対象物の像の可視受信のためのチャネルであり、
前記像は、前記表現と共に以前に記録された前記対象物の像のデータベースであって前記通信チャネルの入力／出力のための前記制御モジュールによりアクセス可能なデータベースと比較される、請求項５に記載のヒューマノイドロボット。
【請求項１１】
第１の通信チャネルは可視メッセージを受信するためのチャネルであり、第２の通信チャネルは音声メッセージを送信するためのチャネルであり、
前記制御モジュールは、前記第１のチャネル上で受信された第１のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第２のチャネル上にその内容が前記信頼水準に依存する少なくとも１つの第２のメッセージを生成することができる、請求項２に記載のヒューマノイドロボット。
【請求項１２】
前記第１のチャネルは、それぞれが期待認識率と関連付けられた表現のリストにより、受信された前記メッセージの像を認識するためのフィルタを含み、
前記第２のメッセージの前記内容は、前記第１のチャネル上の前記第１のメッセージの繰り返しの要求と、前記フィルタの前記表現のサブセットの音声メッセージを受信するための第３のチャネル上で前記対話者により送信される第３のメッセージによる確認の要求と、少なくとも１つの第４のチャネル上の少なくとも１つの他のメッセージの前記対話者による送信の要求と、からなる要求の群から発見的方法により選択される、請求項１１に記載のヒューマノイドロボット。
【請求項１３】
前記チャネルの少なくとも１つは、前記入力と出力のための前記制御モジュールによりマージされた２つのチャネルの出力を入力として受信するハイブリッドチャネルである、請求項１に記載のヒューマノイドロボット。
【請求項１４】
少なくとも１人の対話者とヒューマノイドロボットとの通信の制御方法であって、
異なる様式を使用する通信チャネルによるメッセージ送信の少なくとも２つの工程であってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つの工程と、
前記チャネルの入力／出力を制御する工程と、を含み、
前記制御工程は、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成され、
第１のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される第１のタイプのメッセージを受信する受信チャネルであり、第２のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセー

ジの各タイプのメッセージから選択される前記第１のタイプとは異なる第２のタイプのメッセージを送信する送信チャネルであり、
　前記制御工程は、前記第１のタイプのメッセージを受信する受信チャネル上で受信された第１のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第２のタイプのメッセージを送信する送信チャネル上にその内容が前記信頼水準に依存する少なくとも１つの第２のメッセージを生成することができる、方法。
【請求項１５】
　コンピュータプログラムがコンピュータ上で実行されると請求項１４に記載の方法の実行を可能にするプログラムコード命令を含むコンピュータプログラムであって、
　前記プログラムは、異なる様式に応じた少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される少なくとも２つのチャネルと、前記チャネルの入力／出力の制御のためのサブルーチンと、を含むヒューマノイドロボットを可能にするように構成され、
　前記制御サブルーチンは、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成される、コンピュータプログラム。
【請求項１６】
　異なる様式に応じた少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される少なくとも２つのチャネルと、前記チャネルの入力／出力のための制御モジュールと、を含む少なくとも１つのヒューマノイドロボットと少なくとも１人の対話者との間の通信インターフェースを編集および制御する方法であって、
　前記制御モジュールは、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成され、
　前記選択された機能をプログラムする工程をさらに含み、
第１のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される第１のタイプのメッセージを受信する受信チャネルであり、第２のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される前記第１のタイプとは異なる第２のタイプのメッセージを送信する送信チャネルであり、
　前記制御モジュールは、前記第１のタイプのメッセージを受信する受信チャネル上で受信された第１のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第２のタイプのメッセージを送信する送信チャネル上にその内容が前記信頼水準に依存する少なくとも１つの第２のメッセージを生成することができる、方法。
【請求項１７】
　前記選択された機能をプログラムする前記工程は、音声送信チャネルを装った第１の通信チャネルと、前記少なくとも１人の対話者により前記ロボットの一肢上に加えられた少なくとも１つの身振りを受信するためのチャネルを装った第２の通信チャネルと、を規定する少なくとも１つの下位工程と、
　前記対話者により前記ロボットに伝えられた前記少なくとも１つの身振りと入力との対応を規定する下位工程と、
　前記第１のチャネル上で前記ロボットにより前記対話者に送信される少なくとも１つのメッセージを生成することにより前記入力の仕様を規定する下位工程と、を含む、請求項１６に記載の編集および制御する方法。

【請求項１８】
　第３の触知通信チャネルを定義し、これにより前記対話者が前記第２のチャネルに対し行われた前記入力を検証する下位工程、をさらに含む請求項１７に記載の編集し制御する方法。
【請求項１９】
　前記選択された機能をプログラムする前記工程は、
　音声メッセージを受信するチャネルを装った第１の通信チャネルと音声メッセージを送信するためのチャネルを装った第２の通信チャネルとを規定する少なくとも１つの下位工程と、
　前記第１のチャネル上で受信された第１のメッセージの前記ロボットによる理解の信頼水準を評価するための機能を規定する下位工程と、
　その内容が前記信頼水準に依存する前記第２のチャネル上の少なくとも１つの第２のメッセージの生成を規定する下位工程と、を含む、請求項１６に記載の編集および制御する方法。
【請求項２０】
　コンピュータプログラムがコンピュータ上で実行されると請求項１６に記載の方法を実行するためのプログラムコード命令を含むコンピュータプログラムであって、
　前記プログラムは、異なる様式に応じた、少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つのチャネルと、前記チャネルの入力／出力を制御するためのサブルーチンと、を含むヒューマノイドロボットをプログラムできるように構成され、
　前記コンピュータプログラムは、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択され前記ロボットにより実行される少なくとも１つの機能を制御するための前記サブルーチン内でプログラムするためのモジュールを含む、コンピュータプログラム。
【請求項２１】
　少なくとも１つのパラメータを制御ボックスに渡すことをプログラムするためのモジュールをさらに含む請求項２０に記載のコンピュータプログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
　本発明はヒューマノイドロボットの分野に属する。より正確には、本発明は、ユーザの命令に基づく行為の前記ロボットによる実行と、前記ロボットによる適切な回答の提供と、より一般的なやり方では前記ロボットとその対話者または複数の対話者との「ヒューマノイド関係」の確立と、を可能にするように、この種のロボットとの対話のためのインターフェースをプログラムし採用する方法に適用される。
【背景技術】
【０００２】
　ロボットは、人間の外観および機能：頭、胴、両腕、任意選択的に両手、両脚、両足などのいくつかの属性を有する限りヒューマノイドと称されることがある。外観以外では、ヒューマノイドロボットが達成できる機能は、運動を行い、話し、「推論する」ためのその能力に依存する。ヒューマノイドロボットは四肢または頭を使って歩行し、身振りをすることができる。ヒューマノイドロボットが行える身振りの複雑さは引き続き増加している。
【０００３】
　いくつかのロボットは環境からの刺激に応じて話すことができる。相互作用の可能性を著しく豊かにするいくつかのロボットと人間との対話のための機能を開発できるようにする音声認識および合成ツールの開発もなされてきた。このような言葉を使用する人間－ロ

ボットインタフェースは、特に米国特許出願公開第２００９／２８７６７８号明細書だけでなく米国特許第７，７１１，５６９号明細書により開示されている。

【０００４】

これらの従来技術文書では、音声認識に固有の不完全性は、データベースへのアクセスを必要とする意味論的および／または文脈論的支援に訴えることにより、学習により、認識の疑い（低認識信頼区間、偽陽性、偽陰性など）を解消できるように大きな計算資源を利用することにより、緩和される。これらの手段の利用は、移動運動などのその最重要処理を管理するようにその計算資源が経済的でなければならない多機能ヒューマノイドロボットの場合は、適切でない。

【０００５】

前記ロボット上に埋め込まれた計算資源を使用することにより簡単かつ効果的やり方で、不完全なままとなるセンサとソフトウェアにより行われる認識に関する疑いを解消できるヒューマノイドロボットを有することができれば有利だろう。

【発明の概要】

【課題を解決するための手段】

【０００６】

本発明は、回答の確認の自然なモードを使用するヒューマノイドロボットとの対話のためのインターフェースを提供することによりこの問題を解決する。

【０００７】

この目的のため、本発明は、異なる様式に応じた少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つのチャネルと、２つのチャネルの入力／出力のための制御モジュールと、を含むヒューマノイドロボットを開示する。制御モジュールは、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成される。

【０００８】

有利には、前記通信チャネルは、可聴、可視、触知可能メッセージ、ロボットの少なくとも一部分の移動および／または位置に関するメッセージ、およびデジタルメッセージを送信および／または受信する為の通信チャネルの群から選択される。

【０００９】

有利には、第１の通信チャネルは音声送信チャネルであり、第２の通信チャネルは少なくとも前記１人の対話者によるロボットの少なくとも一部分からの移動および／または位置を受信するためのチャネルである。前記移動および／または位置は、対話者によりロボットに伝えられる入力を表し、入力の前記仕様は、第１のチャネル上で送信されたメッセージによりロボットにより対話者に対し規定される。

【００１０】

有利には、本発明のロボットはさらに、対話者が第２のチャネルに対し行われた入力を検証する第３の触知通信チャネル（ｔａｃｔｉｌｅ　ｃｏｍｍｕｎｉｃａｔｉｏｎ　ｃｈａｎｎｅｌ）を含む。

【００１１】

有利には、第１の通信チャネルは音声メッセージを受信するためのチャネルであり、第２の通信チャネルは音声メッセージを送信するためのチャネルであり、前記制御モジュールは、前記第１のチャネル上で受信された第１のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第２のチャネル上にその内容が前記信頼水準に依存する少なくとも１つの第２のメッセージを生成することができる。

【００１２】

有利には、第１のチャネルは、それぞれが期待認識率と関連付けられた表現のリストに

より、受信されたメッセージを音声認識するためフィルタを含み、第２のメッセージの内容は、第１のチャネル上の第１のメッセージの繰り返しの要求と、フィルタの表現のサブセットの第１のチャネル上で対話者により送信される第３のメッセージによる確認の要求と、少なくとも１つの第３のチャネル上の少なくとも別のメッセージにおける対話者による送信の要求と、からなる要求の群から発見的方法により選択される。

【００１３】

有利には、本発明のロボットは、第１と第２のチャネル上のメッセージの半二重モードでの順番を保証するように、第１のチャネル上での聴取開始信号（ｓｉｇｎａｌ　ｏｆ　ｓｔａｒｔ　ｏｆ　ｌｉｓｔｅｎｉｎｇ）を第２のチャネル上で送信するように構成されている。

10

【００１４】

有利には、前記選択発見的方法は、期待認識率に基づき決定される閾値に関する実認識率（ｒｅａｌ　ｒｅｃｏｇｎｉｔｉｏｎ　ｒａｔｅ）の位置の関数である。

【００１５】

有利には、前記第３のチャネルは、触知的受信のための、またはロボットの一部から移動を受信するためのチャネルである。

【００１６】

有利には、本発明のロボットはさらに、電子メールシステムとインターフェースするためのモジュールを含む。この前記このモジュールは、前記第１と第２のチャネルを使用することにより、前記電子メールシステム上のアカウントのホルダが前記ロボットを、第２のチャネル上で電子メッセージを受信し／読み、第１のチャネル上で電子メッセージを書き／送り、対話により前記アカウントを管理するためのエージェントとして使用できるようにする。

20

【００１７】

有利には、前記第３のチャネルは、第１のチャネルのフィルタの表現のリストに対応する対象物の像の可視受信（ｖｉｓｕａｌ　ｒｅｃｅｐｔｉｏｎ）のためのチャネルであり、前記像は、前記表現と共に以前に記録された前記対象物の像のデータベースであって前記通信チャネルの入力／出力のための前記制御モジュールによりアクセス可能なデータベースと比較される。

【００１８】

30

有利には、第１の通信チャネルは可視メッセージ（ｖｉｓｕａｌ　ｍｅｓｓａｇｅｓ）を受信するためのチャネルであり、第２の通信チャネルは音声メッセージを送信するためのチャネルであり、前記支援制御モジュールは、前記第１のチャネル上で受信される第１のメッセージの前記ロボットによる理解の前記信頼水準を評価し、その内容が信頼水準に依存する前記第２のチャネル上の少なくとも１つの第２のメッセージを生成することができる。

【００１９】

有利には、第１のチャネルは、それぞれが期待認識率と関連付けられた表現のリストにより、受信されたメッセージの像を認識するためのフィルタを含み、前記第２のメッセージの内容は、第１のチャネル上の前記第１のメッセージの繰り返しの要求、フィルタの表現のサブセットの音声メッセージを受信するための第３のチャネル上で対話者により送信される第３のメッセージによる確認の要求、少なくとも１つの第４のチャネル上の少なくとも１つの他のメッセージの対話者による送信の要求を含む、要求の群から発見的方法により選択される。

40

【００２０】

有利には、チャネルの少なくとも１つは、入力と出力のための前記制御モジュールによりマージされた２つのチャネルの出力を入力として受信するハイブリッドチャネルである。

【００２１】

本発明はまた、少なくとも１人の対話者とヒューマノイドロボットとの通信の制御方法

50

を開示する。本方法は、異なる様式を使用する通信チャネルによるメッセージ送信の少なくとも２つの工程であってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つの工程と、チャネルの入力／出力を制御する工程と、を含む。制御工程は、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づきロボットにより受信されたメッセージの理解を向上させるように構成される。

【００２２】

本発明はまた、コンピュータプログラムがコンピュータ上で実行されると本発明の方法の実行を可能にするプログラムコード命令を含むコンピュータプログラムを開示する。本プログラムは、異なる様式に応じた少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つのチャネルと、チャネルの入力／出力の制御のためのサブルーチンと、を含むヒューマノイドロボットを可能にするように構成される。制御サブルーチンは、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づきロボットにより受信されたメッセージの理解を向上させるように構成される。

【００２３】

本発明はまた、異なる様式に応じた少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つのチャネルと、２つのチャネルの入力／出力のための制御モジュールと、を含む少なくとも１つのヒューマノイドロボットと少なくとも１人の対話者との通信インターフェースを編集し制御する方法を開示する。制御モジュールは、第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能とを含む機能の群から選択された少なくとも１つの機能の実行に基づきロボットにより受信されたメッセージの理解を向上させるように構成される。本方法は、選択された機能をプログラムする工程をさらに含む。

【００２４】

有利には、前記選択された機能をプログラムする前記工程は、音声送信チャネルを装った第１の通信チャネルと、前記少なくとも１人の対話者によりロボットの少なくとも一肢の移動を受信するためのチャネルを装った第２の通信チャネルと、を規定する少なくとも１つの下位工程と、対話者によりロボットに伝達される前記移動と入力との対応を規定する下位工程と、第１のチャネル上でロボットにより対話者に送信される少なくとも１つのメッセージを生成することにより前記入力の仕様を規定する下位工程と、を含む。

【００２５】

有利には、本発明の編集及び制御する方法はさらに、第２のチャネルに対し行われた入力を対話者が検証する第３の触知通信チャネルを規定する下位工程を含む。

【００２６】

有利には、前記選択された機能をプログラムする前記工程は、音声メッセージを受信するためのチャネルを装った第１の通信チャネルと音声メッセージを送信するためのチャネルを装った第２の通信チャネルとを規定する少なくとも１つの下位工程と、前記第１のチャネル上で受信された第１のメッセージの前記ロボットによる理解の信頼水準を評価するための機能を規定する下位工程と、その内容が前記信頼水準に依存する前記第２のチャネル上の少なくとも１つの第２のメッセージの生成を規定する下位工程と、を含む。

【００２７】

本発明はまた、コンピュータプログラムがコンピュータ上で実行されると本発明の方法を実行するためのプログラムコード命令を含むコンピュータプログラムを開示する。コン

ピュータプログラムは、異なる様式に応じた少なくとも１人の対話者とのメッセージの自然なコミュニケーションのための少なくとも２つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される２つのチャネルと、チャネルの入力／出力を制御するためのサブルーチンと、を含むヒューマノイドロボットをユーザがプログラムできるように構成される。本コンピュータプログラムは、ロボットにより実行される少なくとも１つの機能であって第１のチャネルと第２のチャネル上で受信／送信されたメッセージを組み合わせる機能と、チャネル上で受信された第１のメッセージに基づき生成された第２のメッセージを送信する機能と、を含む機能の群から選択される少なくとも１つの機能を制御するためのサブルーチン内でプログラムするためのモジュールを含む。

10

【００２８】

有利には、本発明のコンピュータプログラムはさらに、少なくとも１つのパラメータを制御ボックスに渡すことをプログラムするためのモジュールを含む。

【００２９】

本発明のインターフェースはさらに、例えば音声認識が何らかの有効性を持つことができるには周囲騒音が高過ぎる場合に対話が実行される環境に容易に適合化され得る多様式（ｍｕｌｔｉｍｏｄａｌ）確認モードの提供という利点を提示する。したがってユーザには、接触、身振り、または特定の色または形状のディジタル符号の表示により曖昧な回答を置換／確認するようことを勧めることができる。したがってユーザは、ユーザがコンピュータに向かい合った場合またはインテリジェント電話またはタッチパッドを使用する場合に使用し慣れている従来のインターフェースを直観的やり方で置換またはエミュレートできるようにする手段を思うままに有する。

20

【００３０】

さらに、ロボットの表現のモード自体は、対話者の注意を維持するようにおよび提供される回答に関する感情またはヒントを対話者に伝えるように、特にはイントネーション、凝視、身振りを組み合わせることにより、多様式となることができる。さらに、人間同士の自然なコミュニケーションのモードに近づけることにより、本発明のインターフェースは、認識システムの結果を改良すること、「真の仮想性」に埋没されるユーザの経験（すなわち、物理的人間の化身との対話）の質を向上させること、に貢献する。

【００３１】

30

本発明はまた、その設計者により想定されなかったロボットの用途に特に適合化された新しい相互作用シナリオを極めて容易におよびほんの短い時間で作成できるようにするインターフェースを開発するための人間工学的および多目的環境を提供する。

【００３２】

本発明は、いくつかの例示的実施形態の以下の説明とその添付図面からより良く理解され、その様々な特性と利点が浮かび上がることになる。

【図面の簡単な説明】

【００３３】

【図１】本発明のいくつかの実施形態におけるヒューマノイドロボットの物理的アーキテクチャの図である。

40

【図２】本発明の実施形態のいくつかにおける実施に有用なセンサを含むヒューマノイドロボットの頭を示す。

【図３】本発明のいくつかの実施形態におけるロボットの機能の制御を可能にするハイレベルソフトウェアのアーキテクチャの図である。

【図４】本発明のいくつかの実施形態におけるロボットの行動／相互作用を編集しプログラムするための機能アーキテクチャの図である。

【図５】本発明のいくつかの実施形態においてヒューマノイドロボットが受ける応答／刺激のヒューマノイドロボットにより与えられる解釈を改良するために一般的なやり方で適用される処理動作の機能フローチャートである。

【図６】本発明のいくつかの実施形態におけるロボットの行動／相互作用のプログラミン

50

グ論理チャートである。
【図７ａ－ｃ】本発明のいくつかの実施形態における多様式インターフェース（ｍｕｌｔｉｍｏｄａｌ　ｉｎｔｅｒｆａｃｅ）の相互作用の論理的および時間的組み合わせを示すタイムチャートを表す。
【図８ａ－ｅ】本発明の一実施形態における対話言語を変更するための２進選択および選択肢によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。
【図９ａ－ｅ】本発明の一実施形態における対話言語を変更するための選択肢のリストとからの選択によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。
【図１０ａ－ｄ】本発明の一実施形態における選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。
【図１１ａ－ｂ】本発明の一実施形態における選択のリストの選択肢を置換または補足することができるようにし、いくつかの選択肢同士を比較する新しい音声認識試験を実行できるようにする一連の画面を表す。
【図１２ａ－ｄ】本発明の一実施形態における質問のものとは異なる言語で選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。
【図１３ａ－ｄ】本発明の一実施形態における選択のリストのいくつかの選択肢同士を比較する音声認識試験の閾値を検証／修正できるようにする一連の画面を表す。
【発明を実施するための形態】
【００３４】
　図１に、本発明の一実施形態におけるヒューマノイドロボットの物理的アーキテクチャを示す。このようなロボットは、特に２００９年１０月１５日公開の国際公開第２００９／１２４９５１号パンフレットに開示されている。このプラットフォームは、本発明に至る改良のためのベースとして役立った。後の説明では、このヒューマノイドロボットはこの総称用語またはその商標ＮＡＯ（商標）で呼ばれることがあるが、これにより参照の一般性が修正されることはない。
【００３５】
　このロボットは、センサと関節を駆動するアクチュエータとを制御するためのタイプ１１０の電子カードを約２ダース含む。図に示すカード１１０は左足を制御するものである。本アーキテクチャの長所の１つは、関節を制御するカードはその大部分が交換可能であるということである。関節は通常、少なくとも２つの自由度したがって２つのモータを有する。各モータは角度で駆動される。関節はまた、いくつかの位置センサ、特にＭＲＥ（磁気回転符合器：Ｍａｇｎｅｔｉｃ　Ｒｏｔａｒｙ　Ｅｎｃｏｄｅｒｓ）を含む。電子制御カードは市販のマイクロコントローラを含む。これは、例えばＭｉｃｒｏｃｈｉｐ社のＤＳＰＩＣ（商標）であってよい。これはＤＳＰに結合された１６ビットＭＣＵである。このＭＣＵは、１ミリ秒のループスレービングサイクル（ｌｏｏｐｅｄ　ｓｌａｖｉｎｇ　ｃｙｃｌｅ）を有する。ロボットはまた、他のタイプのアクチュエータ、特にその色と強度がロボットの感情を伝えることができるＬＥＤ（発光ダイオード）を含むことができる。後者はまた、他のタイプの位置センサ、特に慣性ユニット、ＦＳＲ（接地圧センサ）等を含むことができる。
【００３６】
　頭１６０は、ロボットの知能、特に、ロボットが割り当てられた使命、特に本発明の枠組み内でゲームへの参加を達成できるようにするハイレベル機能を実行するカード１３０を含む。但し、カード１３０は、ロボット内の他のどこか例えば胴内に置かれてもよいであろう。しかしながらこの場所は、頭が着脱可能な場合、これらのハイレベル機能を交換する、したがって特にロボットの知能としたがってその使命を極めて速やかに完全に変更できるようにすることが分かるであろう。または逆に、同じ人工知能を保持する一方で１つのボデイを別のボデイと交換する（例えば、欠陥のあるボデイを欠陥の無いボデイと）

ことができるようにする。頭はまた、特に言葉または画像を処理するための、または、またＷＡＮ（Ｗｉｄｅ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）広域ネットワーク上の遠隔通信を確立するためにポートを開くために必要なコード化などのサービス入力／出力を処理するための特殊なカードを含むことができる。カード１３０のプロセッサは市販のｘ８６プロセッサであってよい。ＡＭＤ社のＧｅｏｄｅ（商標）（３２ビット、５００ＭＨｚ）などの低消費電力プロセッサが好適なやり方で選択されることになる。カードはまた、一組のＲＡＭとフラッシュメモリを含む。このカードはまた、通常はＷｉＦｉ、ＷｉＭａｘトランスミッションレイヤ上で、随意的にはＶＰＮ内の随意的にカプセル化された標準プロトコルを有するデータの移動体通信用公衆ネットワーク上で、ロボットの外部（行動サーバー（ｂｅｈａｖｉｏｒｓ　ｓｅｒｖｅｒ）、他のロボットなど）との通信を管理する。プロセッサは通常、標準ＯＳにより駆動され、これにより通常のハイレベル言語（Ｃ、Ｃ＋＋、Ｐｙｔｈｏｎ等）、またはハイレベル機能をプログラムするためのＵＲＢＩ（ロボット工学のための特殊なプログラミング言語）などの人工知能のための特定言語を使用できるようにする。カード１２０はロボットの胴内に収容される。これは、カード１３０により計算された命令のカード１１０への送信を保証する計算機が位置する場所である。このカードをロボット内の他のどこかに収容することも可能であろう。しかし胴内の場所は、頭の近くでかつ四肢の十字路に位置し、したがってこれによりこのカード１３０をカード１２０とカード１１０にリンクする接続配置を最小化できるので有利である。このカード１２０の計算機もまた市販のプロセッサである。これは１００ＭＨｚクロックのＡＲＭ９（商標）タイプの３２ビットプロセッサであると有利である。プロセッサの型、オン／オフボタンに近いその中央位置、電源の制御へのリンクは、ロボットの電源（スタンバイモード、非常停止等）を管理するように好適に適合化されたツールとなる。カードはまた、一組のＲＡＭとフラッシュメモリを含む。

【００３７】

３つのレベルを有するこのアーキテクチャは、ロボットが、連携動作と、センサの読み取りと同時にその環境内で発せられた言葉または符号を解釈し、それに反応または応答するような他の行為と、を実行できなければならない本発明の実施形態には特に有利である。

【００３８】

図２ａと図２ｂに、本発明の実施形態のいくつかにおける発明の実施に有用なセンサを含むヒューマノイドロボットの頭の末端図とプロフィール図をそれぞれ表す。

【００３９】

図１の頭１６０は、本発明の実施に有用な感知能力と表現能力をロボットに備えさせるように頭２００ａ、２００ｂとして改善されている。

【００４０】

ＮＡＯは、例えばＫｉｎｇｓｔａｔｅ　Ｅｌｅｃｔｒｏｎｉｃｓ　Ｃｏｒｐにより提供されるＫＥＥＧ１５４０ＰＢＬ－Ａと呼ばれる４つの全指向性マイクロホン２１１ａ、２１２ａ、２１３ａ、２１４ａ、すなわち前部には２１１ａ、裏側には２１４ａ、頭の各側面に２１２ａ、２１３ａを備える（図２ｂも参照されたい）。これらは頭の内部に分散されるので、外部にアクセスするための孔だけが同図では視認可能である。マイクロホンにより行われる音声捕捉に基づき、音声分析および認識システム（例えば、Ａｃａｐｅｌａ（商標）社により提供されるシステムＢａｂＥＡＲ（商標））は、本明細書内の後で提示される適切なインターフェースを有するユーザが自身の用語により豊かにすることができる所定の単語のコーパスを認識する。これらの単語は、ユーザの好みの行動、特にはロボットにより解釈される質問に対する回答をトリガできるようにする。ソフトウェア環境は本明細書の後で示されるようにいくつかの言語を支援する。ＮＡＯはまた、音声の源を検出することができこれにより何人かの対話者間の曖昧さを解消できるようにする。

【００４１】

ＮＡＯは、１秒当たり最大３０個の画像を捕捉することができる２台のＣＭＯＳ　６４０×４８０カメラ（２２０ａ）、例えば０Ｖ７６０と呼ばれるＯｍｎｉｖｉｓｉｏｎ（商

標）ブランドのカメラ（ＣＭＯＳ　１／６ｔｈインチセンサ：３．６μｍの画素）を通して見る。額のレベルに置かれた第１のカメラはその水平線に向けられ、口のレベルに置かれた第２のカメラはその直接環境を調査する。ソフトウェアは、ＮＡＯが見たものの写真および映像ストリームをも読み出すことができるようにする。その環境を感知し解釈するために、ＮＡＯは、ボールとより複雑な対象物とを捜し出すために、対話者を認識できるようにする一組の顔および形状の検出と認識アルゴリズムを搭載する。

【００４２】

ＮＡＯは、その頭蓋の頂部上に置かれた、特にこのアプリケーションのために出願人により例えば３つの区画に分割されて開発された容量センサ（２３０ａ）を備える。特定のアプリケーションでは、４区画以上を設けることができるであろう。したがって、接触により（例えば、アプリケーションにより規定された行動のトリガを可能にする一系列のボタンを押すことにより）、本発明の枠組み内で各ボタンに関連する異なる回答であってよい情報をＮＡＯに与え、提案された選択のリストをスクロールして進み、ヘルプメニュー等にアクセスすることが可能である。システムには、接触があるかどうかを示すＬＥＤが付随する。

【００４３】

ＮＡＯは、例えば明細書の後で説明されるまたはＷｅｂサイトまたはＲＳＳストリームから検索されるモードに応じてプログラムされるその格納スペース内に局所的に存在する任意のテキストファイルを大声で読むことにより自身を表現することができる。頭の各側面上に配置された２つの拡声器２１０ｂを備えることにより、その音声合成システム（例えば、ＡｃａｐｅｌａのＡｃａｐｅｌａ　Ｍｏｂｉｌｉｔｙ）はパラメタータ化可能であり、これにより特に音声の速度および／またはトーンの修正を可能にする。

【００４４】

ＮＡＯに音楽ファイルを送りそれを演奏させることが可能である。ＮＡＯは例えば＿．ｗａｖおよび．ｍｐ３フォーマットを受け入れ、これにより本発明の枠組み内で、音声回答の付属としてまたはその代替として音楽回答または専用音声を提供できるようにする。音楽ファイル以外のフォーマットを受け入れてもよい。

【００４５】

図３は、本発明の一実施形態におけるロボットの機能の制御を可能にするハイレベルソフトウェアのアーキテクチャの図である。

【００４６】

このタイプのソフトウェアアーキテクチャは、特には２００９年１０月１５日公開の国際公開第２００９／１２４９５５号パンフレットに開示されている。このアーキテクチャは、ロボットとＰＣまたは遠隔場所との間の通信を管理するための、そして本発明の実施のために必要なソフトウェアインフラストラクチャを提供するソフトウェアを交換するための、基本機能を含む。このアーキテクチャは、特定のアプリケーションにおいて使用されるソフトウェア機能の特定の言及無しに包括的なやり方で以下に説明される。これらの機能は本発明のロボットの行動を管理するための任意の他のソフトウェア機能として処理されることが分かる。

【００４７】

図３では、例えば移動性理由のために無線リンクにより第１の遠隔端末装置ＴＤ１と通信する第１のヒューマノイドロボットＲＨ１が極めて図式的に表される。遠隔端末装置という表現は、通信ネットワークを手段として、このタイプのヒューマノイドロボットＲＨ１に専用化されたウェブサービスＳＷへのアクセスを提供するサーバープラットフォームＰＦＳから離れた端末を意味するように意図されている。

【００４８】

当然、システムの要素間の通信リンクは有線ベースであってもよく、移動端末は変形形態として携帯電話または携帯型コンピュータであってもよい。

【００４９】

第２のヒューマノイドロボットＲＨ２は、例えばヒューマノイドロボットＲＨ２の移動

性を妨げないように無線リンクによっても第２の遠隔端末装置ＴＤ２と通信する。
【００５０】
遠隔端末装置ＴＤ１、ＴＤ２とサーバープラットフォームＰＦＳは、通信ネットワークＲＣを手段としてネットワーク内でリンクされる。遠隔端末装置ＴＤ１、ＴＤ２だけでなくサーバープラットフォームＰＦＳのウェブサービスのために、またヒューマノイドロボットＲＨ１、ＲＨ２のために、少なくとも１系列の命令を含む少なくとも１つのモジュールに専用化された単一のそれぞれの連結モジュールＢ５、Ｂ２、Ｂ４、Ｂ１、Ｂ３はプロセッサにより実行されるソフトウェア機能を実施する。連結モジュールＢ５、Ｂ２、Ｂ４、Ｂ１、Ｂ３のそれぞれのモジュールＭ５１、Ｍ５２、Ｍ２１、Ｍ２２、Ｍ４１、Ｍ４２、Ｍ１１、Ｍ１２、Ｍ３１、Ｍ３２は、この例では、連結モジュール当たりの数は２で表されるがこの数は連結モジュール毎に異なり任意であってよい。
【００５１】
次に、第１のヒューマノイドロボットＲＨ１を所有する第１の遠隔端末装置ＴＤ１のユーザにより想定されるシステムの動作の全く非限定的な実施例を説明する。ユーザは例えば、第１の遠隔端末装置ＴＤ１上に搭載されたまたは第１の遠隔端末装置ＴＤ１からサーバープラットフォームＰＦＳ上にアクセス可能なソフトウェアアプリケーションによりある数の機能をロボットを介し実行することができる。
【００５２】
例えば、ユーザは、ソフトウェアアプリケーションのグラフィックツールにより、ロボットが１０秒間歩いてから「皆さん、今日は」と言うロボットのアプリケーションを簡単に実行する。このアプリケーションは、例えばモジュール（例えば、モジュールＭ１１）の形式で第１のヒューマノイドロボットＲＨ１内にダウンロードされ、次に第１の遠隔端末装置ＴＤ１を手段としてユーザによりトリガされる。
【００５３】
第１のヒューマノイドロボットＲＨ１は、「歩行」機能を最初に使用しなければならないモジュールＭ１１をトリガする。次に、モジュールＭ１１は、モジュールＭ１１がリンクされる連結モジュールＢ１に要求する接続インターフェース／機能呼び出しモジュールまたはプロキシＰ１を使用する。連結モジュールＢ１は、モジュール内に有する呼び出される機能の場所でネットワーク連結モジュールが要求に応答するまで、それ自身のモジュールに向けられた要求と、反復的やり方でこの動作を繰り返すそれが直接リンクされるネットワークと結合するためのモジュール（子連結モジュール）に向けられた要求と、を行う。この要求に対する応答はまた、プロキシＰ１に直接リンクされた連結モジュールＢ１がこの機能を接続し呼び出す必要があるまで親連結モジュールにより反復的やり方で送信される（逆方向に）。例えば、要求された歩行機能は第２の遠隔端末装置ＴＤ２のモジュールＭ４１内に置かれる。次に、連結モジュールＢ４は、例えばロボットが歩行する継続時間を表す秒単位の整数型の継続時間パラメータと、ロボットの排他的または非排他的歩行（すなわちロボットは歩行しながら別の行為を行うことが許容されるか否か）を表すブーリアン型の排他的パラメータと、を含む「歩行」機能の呼び出しのパラメータを返した。この例では１０秒歩いた後に話すことが望ましいので、この例では歩行機能は、１０秒に等しい継続時間パラメータと１秒に等しい排他的パラメータで呼び出される。
【００５４】
したがって接続インターフェース／呼び出しモジュールＰ１は、所望のパラメータを有する「歩行」機能への接続とその呼び出しとを、あたかもこの機能がローカルに存在するかのように遠隔的に行うことができる。接続インターフェースと機能呼び出しモジュールは、異なる端末またはサーバー上に配置されたモジュールの機能を呼び出すことができる相互通信ソフトウェアを使用する。この機能は、呼び出しモジュールのものとは異なるコンピュータ言語の一系列の命令により書くことができる。プロキシは、例えばＳＯＡＰ相互通信ソフトウェアを使用する。これによりプラットフォーム間および言語間通信アーキテクチャを生成する。
【００５５】

この非局在化（ｄｅｌｏｃａｌｉｚｅｄ）「歩行」機能が行われると、モジュールＭ１１は「話す」機能を呼び出さなければならない。別の接続インターフェースと機能呼び出しモジュールまたはプロキシＰ２は、モジュールＭ１１がリンクされる連結モジュールＢ１に要求をする。連結モジュールＢ１は、当初、格納された一連の命令の形式で実行される機能を手段として、それ自身のモジュールＭ１１とＭ１２に向けられた要求を行い、モジュールＭ１１とＭ１２は例えば、モジュールＭ１２内のこの「話す」機能の有無を返すことになる。連結モジュールＢ１は接続インターフェース／機能呼び出しモジュールＰ２に通知し、次にこのモジュールＰは、ローカル呼び出し型の呼び出しにより、パラメータとして例えば「今日は」と読まれるテキストを有するモジュールＭ１２の「話す」機能を直接呼び出すことができる。このパラメータは連結モジュールＢ１によりプロキシＰ２に送信される。さらに、システムは、ヒューマノイドロボットＲＨ１のこの事例では、外部事象を受信すると前記パラメータの値を更新し前記格納されたパラメータの１つの更新を先の要求に応じてモジュールに通知するように適合化された移動端末の状態を表すパラメータの格納／管理モジュールＳＴＭ（「短期記憶：Ｓｈｏｒｔ　Ｔｅｒｍ　Ｍｅｍｏｒｙ」の呼称）を含む。したがって予め警告を受けたモジュールは、通知されたパラメータの変更の関数として行為を引き受けることができることになる。

【００５６】

先に説明した例に関連して、例えば、格納／管理モジュールＳＴＭはロボットＲＨ１の動き検出器により検出された誰かの外観を表すパラメータの状態を格納することができる。このパラメータがロボットの直接環境内に誰もいないことを表す状態からロボットの直接環境内の誰かの存在を表す状態に移ると、モジュールＭ１１により先に行われた要求に応じて、格納／管理モジュールＳＴＭは事象または信号を介しこの値の変化を予め警告する。次にモジュールＭ１１は、例えば、先に説明した逐次的トリガ（「歩行」と「話す」機能）を自動的にトリガすることができる。

【００５７】

図３の例では、格納／管理モジュールＳＴＭは、遠隔端末装置ＴＤ１の一部を形成するが、変形形態として、他の遠隔端末装置ＴＤ２、サーバープラットフォームＰＦＳ、またはヒューマノイドロボットＲＨ１またはＲＨ２の一部を形成することができる。

【００５８】

格納／管理モジュールのＳＴＭはまた、メモリ内にそれぞれの基準時間間隔にわたるいくつかのパラメータの時間的進展を格納することができる。したがってシステムのモジュールはさらに、ある継続時間からのパラメータの値の進展にアクセスし、これら取るべき行為の進展を考慮することができる。

【００５９】

変形形態として、呼び出される機能のモジュールは、ヒューマノイドロボットＲＨ１、ＲＨ２上の、または通信ネットワークＲＣの遠隔端末装置ＴＤ１、ＴＤ２上のサーバープラットフォームＰＧＳ上に置かれてもよい。

【００６０】

したがって本発明は、機能に対しローカルまたは遠隔呼び出しを行うかどうかに関わらず、ネットワーク全体にプログラムを分散させ移動端末の同一動作を可能にする。

【００６１】

さらに、本アーキテクチャはまた、移動端末の状態を表す一組の格納パラメータを有するようにでき、いくつかの行為を自動的にトリガするようにこの状態の進展を考慮できるようにする。

【００６２】

さらに、格納／管理モジュールはまた、所定時間間隔中にパラメータの値の進展を記録することができ、これによりモジュールがこれらのパラメータの進展のログにアクセスできるようにする。

【００６３】

ロボットのインターフェースを管理するためのＮＡＯＱＩと名付けられたオペレーティ

ングシステムを構成するこれらの通信および格納機能は本発明の実施のために特に有用である。

【００６４】

図４は、本発明の一実施形態におけるロボットの行動を編集しプログラムするための機能アーキテクチャの図である。このようなアーキテクチャについては２０１０年5月２５日出願の欧州特許出願第２０１０／０５７１１１号明細書により説明されている。前記アーキテクチャを実装できるようにするヒューマノイドロボットの行動を編集しプログラムするためのソフトウェアは、Ｃｈｏｒｅｇｒａｐｈｅ（商標）と商業的に命名されており、上記特許文献の一般性を損なうことなくその総称または商標のいずれかで呼ぶことがある。

【００６５】

このアーキテクチャにより制御されるロボットは頭、胴、四肢を有するヒューマノイドロボットであってよく、その各部品は関節で繋がれ、各関節は１つまたは複数のモータにより制御される。このアーキテクチャは、システムのユーザが、仮想ロボット上でシミュレートされ有線ベースまたは無線のリンクによりシステムにリンクされた実ロボット上で実行される行動を生成することによりこのようなロボットを制御できるようにする。

【００６６】

これは、そうするようにプログラムされたコンピュータの画面上の行動（例えば、歩く－真直ぐに、右へ、または左へｎ歩；「今日は」－頭上への片腕の動き；言葉等）と（所与の角度の頭、一肢の）動きとを視認すること、シミュレートすること、実行することを伴う。

【００６７】

図４は、時間的次元を有する事象によりトリガされる制御の表現を例示する処理工程のフローチャートである。事象によりトリガされる制御は、本発明の意味論では「ボックス」または「制御ボックス」４１０により表される。ボックスは、この後定義される以下の要素の１つまたは複数を含むことができるツリー状のプログラミング構造である。

－フレーム４２０の「タイムライン」または時間軸、

－「図表」またはフローチャート４７０、

－スクリプト４９０。

【００６８】

制御ボックスは通常、本明細書の後で詳述されるように、事象情報項目を１つのボックスから別のボックスに通常は送信する接続部により相互にリンクされる。いかなるボックスも、ロボットの行動／動きのシナリオを初期化する「ルートボックス」またはルートに直接または間接的にリンクされる。

【００６９】

フレーム４２０の時間軸は、フレームの前記時間軸が挿入されるボックス内に定義されたロボットの行動と動きとが受ける時間的制約を表す。以下の説明と特許請求範囲では、プログラミングの世界において同じ意味を持って一般的に受け入れられている、タイムラインという用語を使用する。したがってタイムラインはボックスの行動と動きの同期を行う。タイムラインは、毎秒フレーム数（ＦＰＳ：Ｆｒａｍｅｓ　Ｐｅｒ　Ｓｅｃｏｎｄ）の表現で定義された進捗速度に関連付けられたフレームに細切れにされる。各タイムラインのＦＰＳはユーザによりパラメタータ化可能である。デフォルト設定では、ＦＰＳは任意の値（例えば、１５ＦＰＳ）で固定されてもよい。

【００７０】

タイムラインは次のものを含むことができる。

－それぞれが１つまたは複数の行動キーフレーム（Ｂｅｈａｖｉｏｒ　Ｋｅｙ　Ｆｒａｍｅ）または「主行動フレーム（ｍａｉｎ　ｂｅｈａｖｉｏｒ　Ｆｒａｍｅｓ）」４５０を含む１つまたは複数の行動レイヤ（Ｂｅｈａｖｉｏｒ　Ｌａｙｅｒ）４３０であって、上記行動キーフレーム自体が、行動レイヤまたはタイムラインを通過することなく、実際はより高いレベルのボックスにも直接取り付けられることができるボックスの組である１

つまたは複数のチャートまたは「フローチャート」４７０を含むことができる、行動レイヤ４３０、
－それぞれが、１つまたは複数の運動画面４８０を含むことができる１つまたは複数の運動キーフレームまたは「主運動フレーム」４６０を含む、１つまたは複数の運動レイヤ４４０。
【００７１】
行動レイヤは、一組のロボットの行動または主行動フレームを定義する。いくつかの行動レイヤが同一のボックス内に定義されてもよい。これらは次に、ボックスのタイムラインにより同期されたやり方で進行するようにプログラムされる。
【００７２】
行動レイヤは１つまたは複数の主行動フレームを含むことができることになる。主行動フレームは、歩く（「歩行」）、話す（「言う」）、音楽を演奏する（「音楽」）などのロボットの行動を定義する。一定数の行動は、本明細書の後で詳述されるようにライブラリからの単純な「ドラッグアンドドロップ」でユーザにより直接挿入されるように本発明のシステム内に予めプログラムされる。各主行動フレームは、タイムライン内に挿入されるフレームの開始であるトリガ事象により定義される。主行動フレームの終了は、それに続く別の主行動フレームが挿入される限りにおいてだけ、または終了事象が定義される場合に定義される。
【００７３】
運動レイヤは、ロボットの関節のモータの動きをグループ化する１つまたは複数の逐次的主運動フレーム（ｓｕｃｃｅｓｓｉｖｅ　ｍａｉｎ　ｍｏｔｉｏｎ　Ｆｒａｍｅｓ）によりプログラムされる一組のロボットの運動を定義する。実行されるこれらの動きは、本明細書の後で詳述されるアニメーション画面上の行為によりプログラムされてもよい前記モータの到来角度位置（ａｎｇｕｌａｒ　ｐｏｓｉｔｉｏｎｓ　ｏｆ　ａｒｒｉｖａｌ　）により定義される。同一のボックスのすべての主運動フレームはボックスのタイムラインにより同期される。主運動フレームは到来フレーム（ａｒｒｉｖａｌ　Ｆｒａｍｅ）により定義される。開始フレームは、前の主運動フレームの終りのものまたはボックスの開始事象のものである。
【００７４】
主行動フレームと主運動フレームは主行為フレーム（ｍａｉｎ　ａｃｔｉｏｎ　Ｆｒａｍｅ）の一般名で呼ばれる。
【００７５】
同じタイムラインに取り付けられているという条件で、いくつかの主行為フレーム（行動、または運動の）を並列に実行することが可能である。
【００７６】
フローチャートは後で詳述されるように相互に接続された一組のボックスである。ボックスのそれぞれは次に、新しい行動または運動レイヤが取り付けられる他のタイムラインを含むことができる。
【００７７】
スクリプトはロボットにより直接実行可能なプログラムである。本発明の枠組み内では、スクリプトは好ましいやり方でＣ＋＋言語で書かれる。スクリプトを含むボックスはいかなる他の要素も含まない。
【００７８】
ソフトウェアは、Ｗｉｎｄｏｗｓ（商標）、Ｍａｃ（商標）、またはＬｉｎｕｘ（商標）オペレーティングシステムを使用することにより、ＰＣ上またはパーソナルコンピュータ型の別のプラットフォーム上に埋め込まれてもよい。
【００７９】
本発明のヒューマノイドロボットは通常、Ｃｈｏｒｅｇｒａｐｈｅ（商標）ソフトウェアを使用することにより人間と相互作用できるようにプログラムされることになる。この開発アーキテクチャにより可能になる時間的および行動的論理の組み合わせは、本発明の

実施のために特に有利である。本明細書の後で説明される一定数のツールは、本発明の枠組み内での自然な対話のためのインターフェースを有するヒューマノイドロボットを実現するために特に開発された。

【００８０】

図５は、本発明のいくつかの実施形態においてヒューマノイドロボットが受ける応答／刺激のヒューマノイドロボットにより与えられる解釈を改良するために一般的なやり方で適用される処理動作の機能フローチャートである。

【００８１】

長期にわたって、人間は機械と相互作用する多様な手段を開発してきた。これらの手段はテクノロジーの進化に従うので、常により効果的となる。すべての場合において、効果的であるためには、相互作用はそれが何であれプラットフォームとユーザの必要性とに適合化されなければならない。

【００８２】

したがってグラフィックインターフェースとウィンドウ環境は、例えばテキスト領域（テキストボックス）、ＯＫ／キャンセルボタン、チェックマークされるボックス（チェックボックス）、ラジオボタン、または複合ボックス（コンボボックス）などのグラフィックインターフェース部品とも呼ばれる一定数のインターフェース要素（ＧＵＩ要素すなわちグラフィックユーザインターフェース要素）をユーザの自由意思で配置する。グラフィックインターフェースに適合されたこれらの要素は、従来の画面タイプのいかなる視覚的フィードバックも原理的に提供しないヒューマノイドロボット上などでは使用することができない。今や、ロボットとのやりとりは、コンピュータのグラフィックインターフェースとのやりとりと少なくとも同じくらい豊富でなければならない。このとき、人は、ボックスにチェックマークを付け、キーボード上でテキストを入力し、アイコンをダブルクリックし、またはアプリケーションのウィンドウ内の十字をクリックするのと同じやり方で選択肢を選択し、単語を綴る、またはアプリケーションを実行するまたはアプリケーションを離れることができることを望む。人は、人間化され自然なユーザの為のインターフェースを望むからこそこれらの既存要素を単純に複製したいとも思わない。したがって自律ヒューマノイドロボットに適合化されたユーザインターフェース要素を見つける必要がある。

【００８３】

これらの要素はまた、ヒューマノイドロボット行動の作成者にとって容易にパラメタータ化可能でなければならなく、またユーザの言語に容易に適合できるようにしなければならない。既存の自律ロボットは音声認識などの簡単な人間－ロボットインタフェースを導入することができるが、従来技術では、ユーザと開発者のいずれも、地域化される（多言語使用を可能にする）とともに失敗を管理するいかなる多様式（ｍｕｌｔｉｍｏｄａｌ）ユーザインターフェース要素を提供されなかった。実際、今日、多重センサ捕捉および処理能力と移動運動能力とその四肢の多数の自由度とを備えた合理的な大きさと価格のヒューマノイドロボット内に埋め込むことが可能な音声認識のタイプは、ロボット上に搭載可能なコンピュータ資源と電気エネルギー資源により必然的に制限される。これらの資源は実際には、信号の捕捉の安全性と信頼性と動きの実行に必要な制御とを保証できるようにする処理動作に対する優先度により必然的に割り当てられる。したがって音声認識のこの状況における避けられない不完全性の最良の補正を可能にする人間－ロボットインタフェース要素を提供することと、特には、ロボットがユーザから受け取るメッセージに対しロボットにより与えられる解釈についての疑問と収束する対話シーケンスの枠組み内で生ずるロボットにより返される質問とを解決するためのメカニズムによりユーザに良好な相互作用を提供すること、が必要である。

【００８４】

ロボットは人間的基準（ｈｕｍａｎ　ｒｅｆｅｒｅｎｃｅｓ）、すなわち、人間が同じ状況においてとるであろう身振りと行動を返さないので、人間は自然なやり方でロボットに話しかけないということにも注意する必要があるだろう。相互作用はとりわけ、ロボッ

トが人間の方向を見なければ（人間－人間相互作用における習慣的相互作用が無ければ）自然なものとはならない。さらに、人間コミュニケーションとの対比では、多機能ヒューマノイドロボットに搭載されるコンピュータ資源に適合する音声認識のタイプはそれ自体は幾人かのユーザとの相互作用の効果的管理を可能にしない。さらに、たいていのロボットは自然言語をほとんどまたは全く使用しない。音声合成は通常、ロボットが読むことになるロボット用に考案された話または人間により書かれた電子メールであったとしても、人間により予め書かれた語句によりプログラムされる。したがって音声合成は、人間－ロボット相互作用を人間－人間相互作用に可能な限り近付けるようにする要素を欠く。従来技術の人間－ロボットインタフェースは、人間－人間自然相互作用を模擬し、この相互作用の成功に貢献できるようにする十分な多様式（ｍｕｌｔｉ－ｍｏｄａｌｉｔｙ）または相互作用コードを有していない。さらに、ユーザにより既に取得された知識であってユーザが日々使用する知識でもインターフェースが呼び出せば、経験は、はるかに容易となり、ユーザのその部分について少し学習する必要があるだけとなる。したがって、仮想世界の部屋全体にわたって眼を巡らすことが、コンピュータキーボード上の矢印を押すことによるよりも頭を動かすことにより仮想現実ヘルメットによりいっそう本能的に行われることになる。

【００８５】

本発明の解決策は、ソフトウェアとハードウェアを組み合わせた、自律ヒューマノイドロボットに適合化されたユーザインターフェース要素を提案する。上述のように使用される用語である、ＧＵＩ要素をロボットの行動に置き換えることにより、ＢＵＩ要素（ビヘイビアユーザインターフェース要素：Ｂｅｈａｖｉｏｒ　Ｕｓｅｒ　Ｉｎｔｅｒｆａｃｅ　Ｅｌｅｍｅｎｔｓ）、本明細書ではより一般的かつ単純にＵＩＥｌｅｍｅｎｔｓと呼ぶことができる、を定義する。このようなＵＩＥｌｅｍｅｎｔｓは、例えば、次のような行為を簡単なやり方でコード化するように定義されることができる。

－ロボットの頭の３つの触覚センサを同時に叩くことにより即座にアプリケーションを中止する。

－音声認識を使用することによりロボットに尋問する。

－ロボットの触覚センサのうちの１つを叩くことによりアプリケーションの次の工程に移動する。

【００８６】

したがってこれらの単純な要素は、ロボットのすべての行動とアプリケーションに利用可能となるようにまたは所与のプロジェクトの特定の資源を装って生成されるように一般的ライブラリに埋め込んでもよい真の相互作用コードである。

【００８７】

本発明のＵＩＥｌｅｍｅｎｔｓは、行動開発者により使用され容易にパラメータ化することができる要素である。行動をプログラムするための基本ＧＵＩ要素となるのは主にＣｈｏｒｅｇｒａｐｈｅボックスである。特に、これらのボックスのいくつかは、グラフィックインターフェース部品を開発するためのＱｔ（商標）環境により生成されるＷｉｄｇｅｔライブラリを使用することによりＣ＋＋でコード化されたＣｈｏｒｅｇｒａｐｈｅプラグインを含む。

【００８８】

本発明の実施を可能にする機能アーキテクチャの簡略図を図５に表す。

【００８９】

ロボットが対話者とメッセージを交換する通信チャネルの入力／出力の制御モジュール５１０は、図１の中央ユニット１２０内に取り付けられるかまたはそれに関連して取り付けられる。このモジュールは、ロボットが備える特定の通信チャネルの送信／受信前処理手段を物理的または論理的に含む。

【００９０】

これに限定されないが、それぞれが受信チャネルと送信チャネルを有する３つのタイプのメッセージ通信チャネルを同図に表した。

【００９１】

タイプ１の受信チャネル５２１は、人間の聴覚に対応し、ロボットが音声信号（好ましくは意味論的内容を含む音声メール）を取得できるようにする。したがってロボットは図２ａに表されたマイクロホン２１０ａを備えてもよい。このチャネルの出力は通常、音声認識アルゴリズムを実行する特定の信号処理プロセッサにより前処理される。これらのアルゴリズムは、多少複雑であるかもしれなく、そして使用される環境（周囲騒音、複数の話し手など）とおおよそ完全な特定の学習の実施とに応じて変わり得る有効性を有するかもしれない。しかしながらすべての構成において、認識誤りは避けられない。

【００９２】

タイプ１の送信チャネル５３１は、人間の話し言葉に対応しており、ロボットが例えば図２ｂに表された拡声器２１０ｂを手段として話すことができる（すなわち意味論的内容を含む音声メールを読み上げることができる）ようにする。言語、音質、リズム、口調は、状況の関数としてそして感情を表現するために変えられてもよい。しかしながらこれらの音はまた、ビープ音、予め記録された音楽であってよい。例えばモールスシーケンスのビープ音、予め設定されたコードに従う音楽もまた、意味論的内容を有することができると理解される。

【００９３】

タイプ２の受信チャネル５２２は、人間の視覚に対応しており、ロボットがその環境をマッピングできるようにし、アクセス可能なメモリ内に格納されていれば認識できる画像を取得できるようにする。したがって、ロボットは例えば図２ａに表されたＣＭＯＳカメラ２２０ａを備えてもよい。カメラの１つは遠方視野に専用化され、他は近傍視野に専用化されることが好ましい。有利には、画像認識アルゴリズムはロボットの対話者の顔を検出または顔を実際に認識できるように適合化される。繰り返すが、認識性能が何であれ、不確実性または誤りは避けられない。画像認識はまた、その意味がコード化により定義され得るＶＤＵ上でロボットに提示される番号数字または商標などの単純な形に適用することもできる。

【００９４】

タイプ２の送信チャネル５３２は、直接的な人間の等価物の無い人工的チャネルである。このチャネルは、ロボットの体内に埋め込まれたＬＥＤにより生成される光信号の発射を可能にする。数多くのＬＥＤが、特に眼、耳、胴、足上に設けられてもよい。これらは、異なる色を有し、可変周波数明滅能力を備えることができる。このチャネルは、メッセージを送る単純で強力な手段をロボットに備えさせる。特に、特定コードがユーザにより定義されプログラムされてもよい。

【００９５】

タイプ３の受信チャネル５２３は人間らしさと等価なチャネルである。但しこのチャネルはその触覚領域が制限される。触覚領域は例えば、図２ａに表されたセンサ２３０ａなどの触覚センサに集中する。ロボットの対話者は、ロボットに２進（行為の確認）またはさらに複雑なタイプのメッセージを伝えるために触覚センサを作動する。このチャネルにより受信される情報は実際、ユーザにより定義される単一的コード（罰と褒美の意味をそれぞれ有する、舌打ちとなでること）またはモールス型の連続的コードのいずれかに対応することができる。このタイプの通信チャネルを定義するために特定の触覚センサが絶対的に必要ということではない。本明細書のさらに後で説明されるようにメッセージセンサがロボットの腕および／または前腕の位置により表される連続的アナログセンサであって前記位置が対話者によりロボットに伝達される数値を表す同じタイプのチャネルを、それが対話者の接触行為を受信する限り、定義してもよい。実際、いつでも、ロボットはその関節の角度位置を知っており、したがって移動の意味が予め定義されていれば対話者の行為による移動に起因する関節の角度位置の変化をメッセージとして解釈する方法を知っている。一肢（例えば、前腕）の単純な接触もまた、ロボットの関節の角度位置のセンサにより識別することができる。衝撃または持ち上げなどのより急峻な動きは、ロボットの慣性ユニットと足の裏センサ（ＦＳＲ）とによりそれぞれ検出されてもよい。

【００９６】
　タイプ３のタイプの送信チャネル５３３は人間の身振りと等価である。頭は、ヨー角により測定される方位角の変位とピッチ角により測定される仰角の変位との２つの自由度を与えられてもよい。これらの２つの動きは伝統的に承認（ピッチ）または拒絶（ヨー）のメッセージを定義する。これらはまた、ロボットが会話状態にある対話者の方にその視線を向けることができるようにする。肩、肘、手首の関節はピッチ、ロール（ロールまたは右／左ねじれ）、ヨー、ヨーの自由度を備えてもよい。手は開く能力と閉じる能力を備えてもよい。これらの関節の動きの組み合わせにより、このチャネルを通しロボットの対話者に伝達されるメッセージの内容を定義できるようにする。
【００９７】
　他のメッセージ通信チャネル（図には表されない）が存在する、またはそれらが定義されてもよい。特に、ロボットは、赤外線、ブルートゥース、またはＷｉｆｉリンクにより信号を送受信することができる。したがって、特に、この目的のためにプログラムされたリモコン、例えばＡｐｐｌｅ（商標）のｉＰｈｏｎｅ（商標）またはモーションキャプチャおよび／または測位機能を有する別の電話を使用することにより、対話者はこのチャネルを介しロボットにメッセージを送信することが可能である。
【００９８】
　同様に、ロボットはこれらの通信ポートを介し別のロボットにメッセージを送ることができる。
【００９９】
　本発明によると、メッセージ通信チャネルは、異なるタイプのチャネルをハイブリッド型のチャネルにマージすることにより定義されてもよい。したがって音声認識を備えた音声チャネルと画像認識を備えた可視チャネルの出力は、その出力がデータマージ処理により改良されることになる新しいチャネルを生成するように組み合わせられてもよい。このチャネルからの出力に関する出力は、別個に採取された２つの出力より高い信頼水準のアプリオリである。
【０１００】
　ロボットの２人の対話者５４１と５４２もまた図５に表される。当然、ただ１人または３人以上の対話者が本発明の実施シナリオでは可能である。さらに、メッセージの交換に必要な可聴および／または可視信号を送信できるようにするデータリンクによりロボットが置かれた部屋にリンクされるという条件で、対話者はロボットからいくらか離れて置かれてもよい。当然、この場合、物理的接触を必要とするタイプ３の通信チャネルの利用は不可能となる。
【０１０１】
　ロボットのその対話者に対するおよびロボットのその環境に対する相対位置はまた、人間／ロボット対話の性質を特徴付けるようにおよび随意的にその進行を修正するように特定のセンサ（対話者の場所に関連する音声認識；画像認識；超音波センサ等）により測定され、例えば体積、トーン、または表情の解析により解釈され相互参照されることができる。したがって、近づき大声で話す対話者は、ロボットにより脅威と見なされ、関連する身体言語により、または相互作用の実際の修正または遮断により様々な防衛行動をトリガすることがある。
【０１０２】
　これらの様々な通信チャネルの入力／出力の論理的制御はモジュール５１０により行われる。
【０１０３】
　後者は、本明細書のさらに後で説明されるように同じ第１のタイプの送信チャネル上に送信されたメッセージを通して第１のタイプの受信チャネル（例えば、音声チャネル）の入力についての疑問を同一時間に解決できるようにする。上記疑問解決行為は恐らく、同じ第１のタイプのチャネルまたは第２のタイプの受信チャネル（例えば、触覚チャネル）上で対話者によりそれに応じて行われる。第１のタイプのチャネル（例えば、音声チャネ

ル）上で受信されたメッセージについての疑問の解決を要求するメッセージはまた、第２のタイプのチャネル（例えば、ＬＥＤ送信による可視チャネル）上で送信することができ、対話者の疑問解決行為は第３のタイプの受信チャネル（例えば、触覚チャネル）上で行われなければならない。これらの組み合わせは純粋に非限定的例示として与えられ、様々な組み合わせが可能性である。

【０１０４】

通信チャネル５１０の入力／出力の制御モジュールもまた、メッセージの入力をより簡単に組み合わせるために使用することができ、この組み合わせによりロボットの「心」の中の疑問のいかなる可能性も実質的に除去できるようにする。

【０１０５】

受信チャネルにより受信された入力と受信チャネルにより送信された出力とを組み合わせるための機能のプログラミングは、ＢＵＩＥｌｅｍｅｎｔｓを使用することにより単純なやり方で実行されてもよい。

【０１０６】

我々は、選択タイプの制御ボックスまたは選択ボックスからなるＢＵＩＥｌｅｍｅｎｔのタイプについて後で説明するものとする。後者は、閉じたリストから選択を行う方法を表す。それは特に、対話の枠組み内で、限定された数の単語と語句の認識に適合化され、ロボットはユーザの選択を聴く前に質問を投げかけることができる。

【０１０７】

我々は、異なるタイプの選択ボックスとは別個のタイプのＢＵＩＥｌｅｍｅｎｔについて以下に説明する。

【０１０８】

我々は、整数を選択する例によりこの様式を示す。この要素に関し、ロボットは例えばタイプ１の送信チャネル５３１上で、ユーザに利用可能な最小数と最大数について述べ、一方の腕を対話者の方へ伸ばす。ロボットは弱くサーボ制御される。この腕は、図５のタイプ３の受信チャネル５２３を構成することになる。腕の低い位置は最小桁に関連付けられ、高い位置は最大桁に関連付けられる。したがってユーザは桁を選択するためのカーソルとしてロボットの腕を利用する。ロボットは、肩のピッチ関節（ＳｈｏｕｌｄｅｒＰｉｔｃｈ）上で利用可能な触角のおかげで腕の位置を知る。この相互作用を補強するために、ロボットはユーザがロボットの腕を動かしている間にその手を見る。位置の各変化によって、ロボットは選択された桁を言うことができる。ユーザは、タイプ３の別の受信チャネル５２３を使用することによりロボットの頭の中央の触覚センサに触れることにより、選択を検証することができる。特にセンサの精度に対し余りにも多過ぎる数の桁がある場合、一方の腕により粗調整がなされ、そして第２の腕がより正確に選択するようにしてもよい。表情の順序付きリストを数字により表してもよい。このとき上記手順は、ロボットにより告げられたドロップダウンメニューからの選択の様式になる。

【０１０９】

桁を選択できるようにする変形形態は、触覚センサだけを使用することからなる。例えば、

－前方センサを叩くことで、桁のリストにおいて１段階下ることができるようにし、

－後方センサを叩くことで、桁のリストを上ることができるようにし、

－前方または後方センサを押したままにすることで、桁のリスト内のスクロールを加速できるようにし、

－選択は中央センサに触れることにより行われるだろう。

【０１１０】

本発明の使用のシナリオの関数として、可能な組み合わせを大幅に変えることが可能であることが理解される。

【０１１１】

図６は、本発明のいくつかの実施形態のロボットの行動／相互作用のプログラミング論理チャートである。

【０１１２】
　同図により示された例は、ロボットに単語のリストからの選択を提示する対話者とロボットが対話するシナリオである（例えば、なぞなぞのゲームの場合の）。このシナリオでは、タイプ１の受信チャネル、タイプ３の受信チャネル、タイプ１の送信チャネルが使用される。
【０１１３】
　図中のコード６１０により表される行為は、ロボットの対話者の行為：すなわち、例えばロボットにより既に告げられたリストからユーザにより告げられた選択；タイムアウト（すなわち、無選択）；このリスト内の１つまたは複数の単語の理解の確認要求に対する回答「はい／いいえ」である。
【０１１４】
　図中のコード６２０により表される行為は、コード６３０により表される内部変数の状態の関数として起動されるロボットの行為である。これらの内部変数の意味は以下の通りである。
　－ｒ：選択のリストの中からユーザにより告げられた単語のロボットによる認識の確率の比、
　－ｆ：認識失敗の総数、
　－ｔ：タイムアウト（または、所定の時間後の対話者による選択が無い）の数、
　－Ｓ１：認識確率比の閾値１、
　－Ｓ２：認識確率比の閾値２、
　－ｔｍａｘ：可能なタイムアウトの最大数、
　－ｆｍａｘ：可能な失敗の最大数。
【０１１５】
　タイムアウトが処理される一般的な方法は、日々の人間生活の単純原理について投げかけられる問題へ適応「一言も言わない人は誰であろうと．．．ということに同意する」に相当する。
【０１１６】
　同図に表される処理動作の一般的論理について以下に説明する。
【０１１７】
　ＮＡＯはユーザ／対話者に耳を傾け、変数ｆとｔはゼロに初期化される。所定のタイムアウト時間が経過したことを対話者が認めればタイムアウトカウンタはインクリメントされ、タイムアウトの最大数に達すれば相互作用ループが遮断される。
【０１１８】
　このアプリケーションは、開始される際にプレーヤの数を確かめるためにゲームにおいて、ユーザにより行われるロボットの尋問などの特定の行為が、または頭上の触覚センサの１つを押すことにより、このアプリケーションをトリガする決定論的状況において、または人間の存在の検出、時刻、またはより一般的には一日の事象のログ（ロボットにより格納される）などのパラメータの関数としてこのアプリケーションをトリガする人工知能の状況において、のいずれかの状況における行動で初期化されてもよい。例えば、ユーザがこのアプリケーションを呼び出し中であるということを検知すると、ロボットはユーザがそれの何を望むかおよびなぜそれを呼び出したかを知ることができるようにするアプリケーションをトリガする。別の状況では、人間の存在を検出すると、ロボット自体は、ユーザがプレーすることを大いに望み、かつユーザがプレーしてから長い時間経っているゲームを提案するためにアプリケーションをトリガすることができるようになる。
【０１１９】
　対話者がタイムアウトの終了前に選択を告げると、測定された認識確率比ｒは期待認識確率比の閾値Ｓ１、Ｓ２（Ｓ１<Ｓ２）と比較される。これらを判定するやり方については後で説明する。
【０１２０】
　ｒ≦Ｓ１ならば、この単語の認識は失敗と考えられる。失敗カウンタがインクリメント

される。ｆｍａｘに達すれば、単語は未認識であると明確に宣言され、相互作用は遮断される。ｆｍａｘに達しなければ、以下の３つの場合の準備が図に示すようになされてもよい。

－第１番目の失敗（ｆ＝１）では、ロボットは「分かりませんでした」とその対話者に示し、選択のリストの繰り返しからなる機能「ａｃｔｉｖａｔｅＨｅｌｐＷｈｅｎＦａｉｌｕｒｅ」の１つを起動する。

－第２番目の失敗（ｆ＝２）では、ロボットはまた「分かりませんでした」と示し、対話者に選択のリストを提供することと、対話者にその触覚センサを使用するように依頼し、それをどのように使用するかを示すことからなる機能「ａｃｔｉｖａｔｅＨｅｌｐＷｈｅｎＦａｉｌｕｒｅ」の別のものを起動する。

－それを超える（３≦ｆ<ｆｍａｘ）失敗では、ロボットは、効果的な会話の条件が満たされていないということを対話者に示す語句、例えば、通常は前記対話者を促して会話を終了させることになる「雑音が多すぎる」などを告げることができる。

【０１２１】

Ｓ１<ｒ≦Ｓ２であれば、ロボットは実際に聞いたものに関し疑問を抱き、同図に表された手順に従って、認識したと考える単語または表現を発声し、その対話者に「これは正しいですか？」と尋ねることにより疑問を解消する行為を取ることができる。対話者が「はい」と答えるか、またはタイムアウトの終了時に答えなければ、ロボットは回答が正しいと考える。対話者が「いいえ」と答えれば、失敗カウンタはインクリメントされ、ｆｍａｘに達すると、ロボットは、分からなかった、と明確に示し、相互作用は停止する。ｆｍａｘに達しなければ、

－第１回目の失敗（ｆ＝１）では、ロボットは、選択のリストの繰り返しからなる機能「ａｃｔｉｖａｔｅＨｅｌｐＷｈｅｎＦａｉｌｕｒｅ」の１つを起動することができる。

－第２回目の失敗（ｆ＝２）では、その対話者に選択のリストを提供することと、その対話者にその触覚センサを使用するように依頼し、それをどのように使用するかを示すことからなる機能「ａｃｔｉｖａｔｅＨｅｌｐＷｈｅｎＦａｉｌｕｒｅ」の別のものを起動する。

－第３回目からｆｍａｘまでの失敗では、対話者は認識の確率比が改善するまで選択を繰り返さなければならない。

【０１２２】

このようにして、音声認識の不完全性を著しく緩和することと、ロボットとその対話者との会話における流動性を改善することが可能である。

【０１２３】

図７ａ、７ｂ、７ｃは、本発明のいくつかの実施形態における多様式インターフェースの相互作用の論理的および時間的組み合わせを示すタイムチャートを表す。

【０１２４】

これらの図は、図６のチャートに表されたタイプの相互作用をプログラムできるようにする選択ボックスの図である。

【０１２５】

選択ボックスは、項目４１０下で示されたものなどのボックスであるが、これらは自然な対話のための特定の行動の特に効果的なプログラミングを可能にする特定タイプのものである。

【０１２６】

これらの図面内の符号の意味は以下の通りである。

－図７ａでは、

－７１０ａはロボットまたはその対話者の行為／スピーチを示す。

－７２０ａは触覚センサを示す。

－７４０ａは認識ビープ音を示す。

－７５０ａは回転するアニメ化された位置のロボットの顔のＬＥＤを示す。

－７５１ａはロボットの固定位置の顔のＬＥＤを示す。

－７６０ａはロボットの顔のＬＥＤの点滅を示す（受信されたメッセージのロボットによる理解の関数として様々な色であってよい）。
－７７０ａはタイムアウト機能を示す。
－７８０ａは選択ボックスの出力を示す。
－７９０ａは機能「触覚センサメニューに行く」（図７ｂ）を示す。
－７Ａ０は機能「選択のソートに行く」（図７ｃ）を示す。
－７Ｂ０は機能「音声認識メニューに行く」を示す。
－Ｒ１、Ｒ２、Ｒ３は、ロボットが曖昧さ無しに理解した場合、ロボットが理解するがいくつかの疑問を抱く場合、ロボットが全く理解しない場合をそれぞれ示す。
－図７ｃでは、７１０ｃは機能「前ページのメニューに戻る」を示す。　10
【０１２７】
選択ボックス内でプログラムされた処理動作の一般的論理は既に説明したものと同一である。ここで説明される追加要素は以下のとおりである。
－ロボットの顔のＬＥＤｓ　７５０ａの利用、随意的には一問一答を強調するためにＬＥＤ点滅の利用。ＬＥＤは、ロボットがスピーチを検出し解析中であることを示すために固定位置７５１ａにある。
－認識の準備ができている時を示すためにロボットにより発せられる可聴「ビープ音」の利用。実際、処理能力と電源の制限のために、そしてまた認識時のうるささを回避するために、後者は音声合成と同時には活性化されない。したがってロボットにより対話者に投げかけられた質問は後者によりあまりにも早く答えられる必要はない。「ビープ音」は、答え始めるためにピー音を発する。　20
－このやりとりと以前のやりとりの間にこのロボットの履歴ログとユーザのその経験とに応じていくつかの水準のヘルプを利用する可能性。
－プログラミングを容易にするためにいくつかのメニューをナビゲートする可能性。
【０１２８】
次に説明する図面は、上の図４に対する注釈として述べたＣｈｏｒｅｇｒａｐｈｅソフトウェアの選択ボックス部品が、表された例ではタイプ１（音声交換）の受信チャネルと送信チャネルを使用することによりＮＡＯロボットと対話者間の単純または複雑な相互作用をプログラムするために利用される画面ショットである。
【０１２９】　30
図８ａ、８ｂ、８ｃ、８ｄ、８ｅは、本発明の一実施形態における対話言語を変更するための２進選択および選択肢によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。
【０１３０】
図９ａ、９ｂ、９ｃ、９ｄ、９ｅは、本発明の一実施形態における対話言語を変更するためのリストと選択肢からの選択によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。
【０１３１】
図１０ａ、１０ｂ、１０ｃ、１０ｄは、本発明の一実施形態における選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。　40
【０１３２】
図１１ａと図１１ｂは、本発明の一実施形態における、選択のリストの選択肢を置換または補足することができるようにし、いくつかの選択肢同士を比較する新しい音声認識試験を実行できるようにする一連の画面を表す。
【０１３３】
図１２ａ、１２ｂ、１２ｃ、１２ｄは、本発明の一実施形態における、質問のものとは異なる言語で選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。
【０１３４】
図１３ａ、１３ｂ、１３ｃ、１３ｄは、本発明の一実施形態における、選択のリスト　50

のいくつかの選択肢同士を比較する音声認識試験の閾値を検証／修正できるようにする一連の画面を表す。

【０１３５】

通常、選択ボックスは、ユーザが選択の事前定義セットの中から回答を選択できるようにする。選択ボックスは、開発者が直観的かつ可読なやり方で一組の可能な選択を書くことができるようにする表型の部品を呼び出す。開発者が予め知らなければ、選択のリストもまたボックスへの入力として入力することができる。したがって例えば、ユーザのメールを管理するアプリケーションの場合、ロボットはユーザに、別のファイル内に格納されたユーザのアドレス帳から連絡先を選択させることができる。これらのＵＩＥｌｅｍｅｎｔｓは高度にパラメタータ化可能なツールである。したがって音声認識および／または合成を使用するＵＩＥｌｅｍｅｎｔｓは地域化される。例えば選択ボックスは仏語と英語で編集可能である。そのプログラミングのためのグラフィックインターフェースレベルでは、ボックスを編集するための言語を変更するために使用されるＷｉｄｇｅｔ　Ｑｔ（商標）がＣｏｍｂｏＢｏｘであってもよい。

【０１３６】

Ｃｈｏｒｅｇｒａｐｈｅボックスの入力（及び出力）は以下のいくつかのタイプのものであってよい。

１．「バン音（ｂａｎｇ）」：信号が送られる。

２．数字：入力は整数または浮動小数点の数を取り出す。

３．文字列：入力は文字列を取り出す。

４．動的。

【０１３７】

動的型の入力（各々出力）はＡＬＶａｌｕｅを取り出す（各々出力する）。

【０１３８】

ＡＬＶａｌｕｅｓはＮＡＯＱＩライブラリ内に記載された一般型（特には、整数、浮動小数点、配列、ブーリアン、文字列、未初期化ＡＬＶａｌｕｅである「ｂａｎｇ」）の集合である。動的型の入力はアプリケーションの展開を極めて柔軟なやり方で管理する。特に、様式間および／または様式内確認モードの選択と支援の提示は、可能な選択の数に応じてそれらを起動するためにロボットの対話者に提供される。

【０１３９】

したがって、この入力が動的型であるという条件で、ｐｙｔｈｏｎ表（配列型）をＣｈｏｒｅｇｒａｐｈｅボックスに対する入力として入力することが可能である。

【０１４０】

本発明を実施するために使用されるＣｈｏｒｅｇｒａｐｈｅソフトウェアは、ブーリアン型のボックス（チェックボックス）、文字列タイプのボックス（テキストボックス）、編集可能またはエンドユーザによらない文字列タイプの多肢選択のボックス（コンボボックス）、整数または浮動小数点浮動小数点型のボックス（スライダ）、または他のタイプのボックスのパラメータを含む。例えば、その行動またはアプリケーションにおいて選択ボックスを使用するプログラマは、ブーリアンパラメータ「検証された選択を繰り返す」（仏語では、「Ｒｅｐｅｔｅｒ　ｌｅ　ｃｈｏｉｘ　ｖａｌｉｄｅ」）のチェックマークを付けるまたはチェックマークを外す可能性を有する。これは、ＮＡＯがユーザにより検証された選択を系統的に繰り返すかどうかを定義するので、相互作用中のＮＡＯの行動に影響を及ぼすことになる。

【０１４１】

音声認識の欠陥を緩和するために、診断ツールは音声相互作用の成功を最大化できるようにする。したがって、選択ボックスにおいて、開発者が表中の単語のリストの書き込みを完了すると、これらの単語の百分率認識（例えばロボットにより確かに認識される単語に対応する１００％、ロボットが認識しない単語に対応する０％）を示すことになるこのツールを実行することができる。この診断は、音声合成により話された単語（ユーザが話すものに近いと仮定する）と音声認識により予想される単語との比較により行われる。さ

らに、選択毎に相互作用の機会と自然さとを最大化するように、いくつかの表現を定義することができる。したがって、メールを送信するようにロボットに依頼するために、開発者は、ユーザの思うままに「メールを送る」、「メッセージを送る」、「電子メールを送る」などのいくつかの語句を配置することができるようになる。ユーザは、その後、最終的に同じことを言うようにこれらの様々な表現同士間の選択肢を有することになる。

【０１４２】

本発明の解決策はまた、幾人かのユーザの存在を管理しない音声認識の問題を解決できるようにする。人間は、数人と話をする際にはコミュニケーションが困難であると認識しているので、一人ずつ話すことにより適応する。この状況は、よく知られたフランスの「ｔｕ」形式のロボットによる利用などの明確に単一ユーザ相互作用コードの存在により容易にされる。

【０１４３】

欠点のある音声認識は、人間－ロボットインタフェースが特に失敗の状況を最もよく管理し、ユーザに正しい時に話させ（これは相互作用コードに関わる）、対話の代替であってより効果的な解決策を利用できるようにさせなければならないということを必要とする。

【０１４４】

本発明の枠組み内では、音声診断機能がこのタイプの問題を解決できるようにする。この機能は、試験対象の単語を音声合成ソフトウェアである、ｔｅｘｔ－ｔｏ－ｓｐｅｅｃｈにより発音させることにより実行する。次に、この試験対象の単語は音声認識により解析される。より正確には、同じ単語は例えば３回発音され、毎回、単語を発音するやり方の代表的サンプルを有するように音声とそのピッチの速度を変化させる。次に、音声認識により返された３つの認識率が平均化される。単語の推定百分率認識はこの値である。音声診断の２つの可能なモードがある。

【０１４５】

その「一斉（Ｔｏｇｅｔｈｅｒ）」モードは次のように動作する。選択ボックス内に記録されたすべての単語は音声認識により聴取され、次にＮＡＯが、他で説明されるように推定認識率を計算する。

【０１４６】

「一つずつ（Ｏｎｅ　ｂｙ　Ｏｎｅ）」モードは次のように動作する。他のライン上の他の可能な選択肢だけでなく所与のラインに関し、解析対象の単語が音声認識により聴取されるが、それと同じライン上に置かれたその別の単語は聴取されない。この診断の利点は、２つの「同義語」が互いに似ている場合、例えば「ｃｏｕｃｏｕ！」と「ｃｏｕｃｏｕ　ｔｏｉ！」、推定認識率は「一斉」モード（推定認識率は、これらが音声認識によりしばしば混同されるので極めて悪くなるであろう）ほど低くはならない。実際、２つの同義語がロボットにより混同されたとしても深刻ではない。

【０１４７】

診断が各ラインに対して行われると、同義語が推定認識率の降順で配置され、最良の同義語の認識率はラインの終わりに記録される。

【０１４８】

したがって選択ボックスは、ロボットがそれを正しく認識または解釈したかが定かでないときにユーザに回答を確認することを依頼するようにプログラムされる。このメカニズムは、聴力障害を有する人間または理解を困難にする環境に陥った人間により使用されるものと同一である。ロボットは、ユーザの回答の理解のレベルに応じて異なる反応を有することになる。次に、いくつかの閾値（例えば、図５に対する注釈として定義された閾値Ｓ１とＳ２）は、認識ソフトウェアにより計算される認識信頼度の関数として固定される。例えば、第１の認識閾値Ｓ１に達しないとき、ロボットはプレーヤに回答を繰り返すように依頼し、第１の閾値Ｓ１に達したが第２の高い認識閾値Ｓ２に達しないとき、ロボットは、それに対する回答が疑問を解消できるようにする質問を投げかけることになる。ロボットはまた、ユーザがロボットに対し正しく答えるようにヘルプを与えることができる

。すなわちロボットは可能な選択のリストを与え、それとの相互作用手段を示すことができ、投げかけられた質問があればこれを繰り返すことができる。相互作用コードはまた、音声認識の欠陥を緩和するのに極めて役立つ。実際、音声認識は、ロボットが話している間はロボットに話しかけることを可能にせず、音声認識の実行とそれが実際に活性化される瞬間との間の遅延はかなり長い。したがって、音声認識が実行されると可聴コードが働き、話すことができるということをユーザに示す。その後、かなり直観的な可視コードと回転する耳のＬＥＤが、ユーザに、ロボットは聴いているということを知らせる。音声認識を使用するＵＩＥｌｅｍｅｎｔｓはまた、この音声認識に代替手段を提供し、繰り返される理解の問題の場合（これは例えば極めて雑音の多い環境によるかもしれない）にもユーザが成功裡にコミュニケーションできるようにする。これらの代替手段は触知可能、可聴、可視的なものであってよい。例えば、選択ボックスはユーザが触覚センサを使用することにより回答を選択できるようにし、前方センサを押すことで選択のリスト内を進むことができるようにし（このとき、ロボットは各選択を告げる）、後方センサはこのリスト内を後退できるようにし、中央センサは選択を検証できるようにする。ロボットが様々な選択を告げることと、ユーザが検証したい選択を聞いたときに「ＯＫ」と言うこと、とを構想することも可能である。または、そうでなければ、確認のために、「はい」または「いいえ」と答える代わりにユーザはロボットの一方の腕を押すことができる。図５に対する注釈として定義された様々なタイプ１、２、３の通信チャネルの入力／出力の制御モジュールは、選択ボックスの様々な入力／出力間のリンクを通してこれらの組み合わせを管理するための機能を単純かつユーザフレンドリなやり方で生成できるようにする。

【０１４９】
　通常、本発明の解決策は、インターフェースの人間化、すなわち人間－人間インターフェースの模擬形態を提案する。我々は、スピーチ（もちろん、すなわち発言された単語群）だけでなく口調と視覚的要素との３つの主要素が２人の人間同士の直接コミュニケーション中に作用し始めるということを知っている。その証しとして、それらの進展を通して、書くことまたはインスタントメッセージなどの間接コミュニケーション手段を観測することにより、対話中の情報不足が一般的な法則として、直接コミュニケーションの代用物と句読点または最近の顔文字などの代用物の追加によりどのようにして緩和されることができるということを極めて明確に理解することが可能である。すべての場合において、今日の大きな技術的進歩にもかかわらず、これらの基本要素が全体として人間－ロボットコミュニケーションのために置き換えられることは依然として難しい。但し、対話の演出を改良する人工的代用物を見つけることは可能である。ロボットの音声合成と音声認識はスピーチの等価物を可能にする。したがってこれらは人間とのコミュニケーションの柱である。ヒューマノイドロボットはさらに、対話の視覚的要素の大部分、すなわち身振りと顔の表情、を表現することができるという利点を有する。実際、その擬人的ボディにより、その移動は車輪上のロボットほど容易でないが、その身振りは人間行動に、より簡単に基づくことができるので、人間の動きとして容易に解読することができる。このとき、コミュニケーションはより自然に実行される。

【０１５０】
　それにもかかわらず口調と顔の表情は、固定された顔と口調とを有するロボットには欠けている。但し、これらの２つの要素は、これらの要素を伝える他の機能とコードにより補償される。これらは、ユーザによる多少長い学習を必要とする。このとき、目的は、この学習を可能な限り短くし、したがってユーザが既に知っているものに対し可能な限り一貫性がありそれに近いコードにすることである。

【０１５１】
　Ｂｅｎ　Ｓｈｎｅｉｄｅｒｍａｎの著作である、「Ｄｅｓｉｇｎｉｎｇ　ｔｈｅ　Ｕｓｅｒ　Ｉｎｔｅｒｆａｃｅ：Ｓｔｒａｔｅｇｉｅｓ　ｆｏｒ　Ｅｆｆｅｃｔｉｖｅ　Ｈｕｍａｎ－Ｃｏｍｐｕｔｅｒ　Ｉｎｔｅｒａｃｔｉｏｎ」（１９９７年発行：ｈｔｔｐ：／／ｗｗｗ．ｃｓ．ｕｍｄ．ｅｄｕ／ｈｃｉｌ／ｐｕｂｓ／ｂｏｏｋｓ／ｄｔｕｉ．ｓｈｔｍｌ）の中で述べられ、通常はグラフィックインターフェースに適用される人間工学の基

本法則を適合化することにより、単純でかつ一貫性のあるコードとしたがって自然でかつ流動的相互作用を実現する。これらの法則は次の原理：コードとインターフェース要素との一貫性、上級ユーザのためのショートカットの存在、行われる行為に関する即時復帰の存在、対話の明確な終了、誤りの簡単な管理、フィードバックの可能性、ユーザは相互作用中は自分自身をマスターと感じるにちがいないこと、そして最後に、ユーザの短期記憶の軽い刺激について述べている。

【０１５２】

音声認識および合成は、特に自然言語の欠落と限定数の単語だけを認識できるようにする単独ユーザ認識とにより、限定的である。本発明の解決策は、十分に自然な人間－ロボット相互作用を提供するように、ロボットによる自然言語の不使用の問題を解決する。既に、最良の使用はロボットの音声合成からなる。特に、音声合成および／または認識を使用するロボットのＵＩＥｌｅｍｅｎｔｓのほとんどは地域化される。したがって仏語を話す（英語を話す）ユーザは仏語（英語）でロボットと対話することができこれにより相互作用の成功を最大化する。その後、最良の使用は、ロボットの反応性を改良するとともに人間－ロボットコミュニケーションの成功を容易にするために、タイミングと相互作用コードからなる。したがって、選択ボックスは、ユーザからの回答の待ち時間のようないくつかのパラメータを提案する。したがってユーザが全く答えなかったということを考慮するまでにロボットがあまり長く待たないということだけでなく、音声認識を正しい時に起動することが可能となるためにロボットが十分に長く待つことも保証される。相互作用コードは、身振り、可聴および／または可視のものであってよい。したがって音声認識の終了の可聴ビープ音は、ユーザに、ロボットがもはや聴いていないということを知らせる。

【０１５３】

さらに、本発明の解決策では、コミュニケーションは、異なる様式のいくつかの通信チャネルの利用とロボットの一部分に関する特定の行動とによりより自然なものにされる。したがって、音の場所と顔の検出（特にその位置）との利用により、ロボットが人間の対話者の方に頭を回転できるようにする。これは別の人間に対処する際の確定事実であるように思われる。ロボットはまた、特に名前、固有の特徴（例えば、ロボットにより行われた会話と行動の履歴ログ）を利用することにより人間に対処するように、対話者識別（顔認識、音質、声紋など）を実施することができる。ロボットはまた、触覚センサをなでたかどうかに応じて行動をユーザが何と考えた（その人はその行動が好んだ）かを知ることができ、その後、例えば言語コミュニケーション中にその行動を実行することを提案する。ロボットは、状況に適合したやり方で行動しようとする。したがってロボットはアニメーションを再生し、そのＬＥＤを使用し、音を放送することができ、これにより人間が話すときに人間がなす本能的な身振り（両手などにより話すこと）を模擬できるようにする。ロボットはまた、うなずくことができる。いくつかの検討、特にＪｕｓｔｉｎｅ　Ｃａｓｓｅｌｌの記事、「Ｓｏｃｉａｌ　Ｄｉａｌｏｇｕｅ　Ｗｉｔｈ　Ｅｍｂｏｄｉｅｄ　Ｃｏｎｖｅｒｓａｔｉｏｎａｌ　Ａｇｅｎｔｓ」（２００５年発行：ｈｔｔｐ：／／ｃｉｔｅｓｅｅｒｘ．ｉｓｔ．ｐｓｕ．ｅｄｕ／ｖｉｅｗｄｏｃ／ｄｏｗｎｌｏａｄ?ｄｏｉ＝１０．１．１．１２４．９８５３&ｒｅｐ＝ｒｅｐ１&ｔｙｐｅ＝ｐｄｆ）においてなされた検討は、人々は、うなずきが無作為に生じた場合でも機械または化身がうなずくと機械または化身が会話にさらに興味を持っているようだと考える、ということを証明できるようにした。これらのあらゆる身振りの信号（頭、腕または手により同意または反対することなどの）、音声信号、相互作用の年表（ｃｈｒｏｎｏｌｏｇｙ）、対話者の場所、またはそうでなければユーザの意図（ユーザは前進しているのか後退しているのか）の検出は、人間－ロボット相互作用を通常の人間の規範により近付けることによりそれをより自然かつ効果的にすることができる。したがってこれらはまた、制限的音声認識に関する問題のいくつかを解決する。これらは本発明の利点の１つである。図８～図１３の画面ショットは、ＮＡＯロボットと対話者との対話のいくつかのケースをプログラムしてもよいやり方を説明する。

【０１５４】

行動において選択ボックスを使用するためには、選択ボックスをＣｈｏｒｅｇｒａｐｈｅのデフォルトライブラリから図（図８ａ）内にドラッグアンドドロップする必要がある。質問を仏語と英語で編集できるようにするローカライズテキストボックスが追加される。質問「あなたのお気に入りの動物は何ですか？」はローカライズテキストボックスのテキストプラグイン内に英語で書き込まれる（図８ｂ）。人はまた、質問を仏語で編集することを望む。したがって、ボックスのプラグインのコンボボックスが用いられ、仏語が選択される（図８ｃ）。このとき、質問のテキストは、編集されないときは空であるテキストプラグイン内に仏語で書き込まれる："Ｑｕｅｌ　ｅｓｔ　ｔｏｎ　ａｎｉｍａｌ　ｐｒｅｆｅｒｅ？"（図８ｄ）。質問は、ロボットにより正しく管理されるように選択ボックスに入力されなければならない。ローカライズテキストボックスの出力は選択ボックスのｏｎＳｔａｒｔ入力にリンクされる（図８ｅ）。我々は今、選択ボックスのコンボボックスにより示されるように選択を英語で編集するものとする。デフォルトの選択がボックスから取り除かれる。第１番目の選択「センザンコウ」が表の第１行に書き込まれる（図９ａ）。第２行目では、別の動物「蜘蛛」が提案されるが、同義語「タランチュラ」がスラッシュ「／」により蜘蛛から区切られて追加される。これは、ロボットが蜘蛛とタランチュラが同意語であると考えることを意味する（図９ｂ）。選択の編集は例えば「兎」と「ポニー」を追加することにより終了される。行の数は選択が追加される際に自動的に適応することに留意されたい（図９ｃ）。コンボボックスは選択ボックスを編集するための言語を仏語に切り替えるために使用される（図９ｄ）。英語と全く同様に、選択のリストが書き込まれ、これにより「ｐａｎｇｏｌｉｎ」、「ａｒａｉｇｎｅｅ／ｔａｒｅｎｔｕｌｅ」、「ｌａｐｉｎ」、「ｐｏｎｅｙ」を得る（図９ｄ）。

【０１５５】

しかしながら、我々はロボットが実際これらの選択を認識するかどうかを知らない。次に、我々は音声診断機能を実行したい。我々は「プラス」をクリックする（図１０ａ）。次に、我々は評価アイコンをクリックする（図１０ｂ）。行毎に、単語は評価される（図１０ｃ）。同義語の場合、最良のものが初めに配置され、示される百分率はこのときの最良同義語の結果である。音声診断は終了し、このとき「ｐｏｎｅｙ」は極めて不完全に認識されている危険があることが分かる（図１０ｄ）。

【０１５６】

次に、我々は、単語を変え、「ｃｈｅｖａｌ」をその代りに入れることを決定する（図１１ａ）。診断が再実行される。「Ｃｈｅｖａｌ」は８２％の優れた得点を獲得し、我々はそれを保持する（図１１ｂ）。

【０１５７】

我々は英語に切り替え、単語に対し診断を英語で実行する（図１２ａ、１２ｂ）。次に、「ｃｈｅｖａｌ」の翻訳として同義語「馬」が「ポニー」に加えられる（図１２ｃ）。診断が再実行され、「ポニー」より良い得点を有する「馬」が第１番目の位置に自動的に配置されることに留意されたい（図１２ｄ）。

【０１５８】

我々は、次に調整してもよいパラメータを編集する。我々は選択ボックスの左下のサムホイールキーをクリックする（図１３ａ）。パラメータ化ウィンドウが開く（図１３ｂ）。我々はブーリアンパラメータ「腕を動かす」にチェックマークを付ける（図１３ｃ）。したがってロボットは話している間、その腕を動かすことになる。我々は、この新しいパラメータ化を検証するためにＯＫをクリックする。

【０１５９】

行動の一般的入力はローカライズテキストボックスの入力にリンクされ、選択ボックスの出力は行動の一般的出力にリンクされる（図１３ｄ）。

【０１６０】

上に述べた例と同様にプログラムされたソフトウェアの例示的な動作について次に説明する。

【０１６１】

ロボットは、Ｃｈｏｒｅｇｒａｐｈｅアイコン「すべてのモータのオン／オフをスレーブ化する」によってスレーブ化され、次に、姿勢ライブラリの「ｉｎｉｔ　ｐｏｓｅ」位置によって直立状態にされる。ロボットの言語は、そのウェブページ上に存在するパラメータによって仏語に設定される。行動は、Ｃｈｏｒｅｇｒａｐｈｅのプレーアイコンによってロボット上で開始される。

【０１６２】

腕を動かしながら、ロボットは「あなたのお気に入りの動物は何ですか？」と尋ね、次に、傾聴音声信号（ｌｉｓｔｅｎｉｎｇ　ｓｏｕｎｄ　ｓｉｇｎａｌ）を発する。聴いている間、眼は耳と同様に青色になり、頭の触覚センサは青色で明滅する。

【０１６３】

次に、ユーザは「イルカ」と答える。何が言われたかを解析している間、ＮＡＯの眼は黄色になる。ＮＡＯは回答を理解しなく、眼は２度赤色で点滅し、耳は２度青色で点滅する。ＮＡＯは、腕を動かしながら「分かりませんでした。あなたが回答できます：センザンコウ、蜘蛛、兎または馬。あなたのお気に入りの動物は何ですか？」と言い、聴取段階（ｌｉｓｔｅｎｉｎｇ　ｐｈａｓｅ）に戻る。

【０１６４】

次に、ユーザは「兎」と答える。ロボットは確信が持てないが、センザンコウと理解したと信じる。眼が緑色で一回点滅する。次に、ロボットは活発に腕を動かしながら、「私はセンザンコウと理解しましたが、正しいですか？」と言う。ユーザは「いいえ」と答える。ロボットは眼を赤色で一回点滅し、そしてその腕を動かしながら支援を開始する。「センザンコウ、蜘蛛、兎または馬？あなたは、私の触覚センサの助けを借りて回答を選択することもできます。あなたのお気に入りの動物は何ですか？」そして傾聴モード（ｌｉｓｔｅｎｉｎｇ　ｍｏｄｅ）に戻る。次に、ユーザは前方触覚センサを押し、ロボットはその眼を青色で一回点滅し、「センザンコウ」と言う。次に、ユーザは再度押し、ロボットはその眼を青色で点滅しながら「蜘蛛」と答える。３回目に、ロボットは眼を青色で一回点滅ながら「兎」と言う。次に、ユーザは、ロボットの選択を検証するために中央の触覚センサを押す。ロボットは眼を一度緑色で点滅し、次に「兎」と繰り返し、ボックスと行動を終了する。

【０１６５】

以下に説明されるものなどのような、ロボットの通信チャネル間の他の相互作用が可能である。

【０１６６】

選択ボックスは、好ましいやり方で、ユーザの選択を認識するように触覚センサと組み合わせて音声認識を利用する。別の可能性はロボットの視覚（特に画像認識）を利用することである。これは概念認識ではなく物体認識である。瓶が示されれば、画像認識は別の商標のものではなくこの同じ瓶を認識することになる。本発明を実施できるようにするバージョンの開発ソフトウェアの可能性の１つは、このソフトウェア内にロボットのカメラリターン（ｃａｍｅｒａ　ｒｅｔｕｒｎ）を有することである。ユーザはロボットに物体を示し、Ｃｈｏｒｅｇｒａｐｈｅにおいて得られた画像を見て、画像内の興味のある対象物を意のままに特定することができる。ユーザはそれに名前を付ける。次に、ロボットは対象物を解析し、それを画像のデータベースに格納する。次に、ユーザは選択ボックスの可能な選択としてこれらの画像を利用することができる。例えば、ユーザが選択ボックスに「瓶」、「コップ」、「雑誌」などの対象物の名前を記入することを望む場合。ユーザは選択ボックスにこれらの単語を記入し、次に瓶、お気に入りのコップ、雑誌のカバーを取り上げ、それらをロボットが先に説明したように解析するようにロボットに示す。次に、選択ボックスはロボットの画像のデータベース全体を探索する。「コップ」と称する対象物が存在すれば、ＮＡＯはユーザの話を聴くと同時にそれを探す、等々、他の単語も同様である。こうして、ユーザは、ユーザの選択を聴くＮＡＯ上のこのボックスを開始する。ユーザは「瓶」と言ったがロボットは理解しない。２人のやり取りの後、ロボットは、「瓶」、「コップ」、「雑誌」がそのデータベース内にあるのでユーザにこれらを示すこ

とができることを説明する。ユーザは、ロボットの話を聴きながら、記録に役立つ（または同じ商標の）瓶を示すことができる。次に、ロボットはあたかも単語「瓶」を認識していたかのように振る舞う。

【０１６７】

本発明の枠組み内では、ロボットのユーザのメッセージングアカウントを受信する／読む、書く／送信する、管理するエージェントとして機能するようにロボットをプログラムすることも可能である。このアプリケーションについて以下に説明する。

【０１６８】

メールアプリケーションにより、ＮＡＯは特に電子メールを読む、電子メールに応答する、または連絡先に電子メールを送信するだけでなく、受信されたメールの筆者を連絡先に加える、メッセージを削除する、メッセージを未読としてマーキングする、再読する、次のまたは前のメッセージを読むこともできる。

【０１６９】

このアプリケーションでは３つの選択ボックスが使用され、これを不可欠の要素とする。単語は音声診断によって選択されている。

【０１７０】

アプリケーションが実行されると、ロボットはユーザが新しいメッセージを受信したかどうかを調べることから始める。そうならば、ロボットは最初の新しいメッセージを読み、次に、質問を伴わずに選択ボックスを実行する。そうでなければ、ロボットは、この同じ選択ボックスを次の質問を伴って実行する：「あなたは私に何をして欲しいのですか？」。したがって、質問を伴うまたは伴わずに選択ボックスを実行できるということが、メールアプリケーションにおいて利用される。この選択ボックスは、ユーザがＮＡＯの可能な行為の中から選択できるようにする。これらの行為はボックスのプラグインの表に書き込まれる。「タイムアウト」選択ボックス出力は、タイムアウトの場合にＮＡＯが次のメッセージを読むので、役立つ。このとき、パラメータ「無返答時の最大繰り返し回数」は１に設定される。ロボットは最初のタイムアウト時にこの選択ボックスから出る。さらに、ユーザによる選択後にロボットは理解したものを明確に示す特定のアニメーションまたは行為を実行するので、パラメータ「検証された選択を繰り返す」は非活性化される。ブーリアンパラメータ「頭を動かす」、「腕を動かす」、および「脚を動かす」のおかげで、ロボットはその決定に合わせたアニメーションにより活気付けられることになる。

【０１７１】

例えば、このボックスの可能な選択は以下のものである。

－返答を記録する／このメールに返答する／そのメールに返答する／返答する。

－再度読む／このメールを再読する／そのメールを再読する／再読する。

－次／次を読む／次のメールを読む。

－前の／前のメールを読む／前のメッセージを読む。

－未読としてマーキングする／保管する／後で再読する。

－削除する／メールを削除する／メッセージを削除する。

－メールを書く／メールを送る／送る。

－連絡先に加える。

－出る／やめる／飛ばす／停止／止める／キャンセルする／静かにする。「出る」は選択ボックスのデフォルト選択の１つであり、ここではメールアプリケーションから出られるようにする。

【０１７２】

ユーザが「メールを書く」選択肢を選択すれば、ユーザはまずアドレス帳から連絡先を選択しなければならない。この選択機能を行うために、「あなたは誰宛に書きたいのですか？」という質問を伴った選択ボックスが入力として使用される。選択のリストは可変である。その結果、選択ボックスの表は記入されなく、連絡先リストがそれを保存するファイルから読み出され、動的タイプの選択ボックスの「ｃｈｏｉｃｅｓＬｉｓｔ」入力に送られる。この時、パラメータ「検証された選択を繰り返す」は、Ｎａｏがメッセージを誰

宛に送るかを実際に理解したということをユーザに示すために活性化される。
【０１７３】
タイムアウトの場合に誰宛にもメールを送らず実際にはメールの送信をキャンセルしメインメニューに戻ることができるように、パラメータ「無返答時の最大繰り返し回数」は例えばそのデフォルト値３に設定される。同様に、アプリケーションのデフォルト選択である「出る」と言うことにより主メニューに戻ることができるようにするヘルプ機能はユーザが連絡先をもはや思い出さない場合のためのものである。この場合、例えば触覚センサによりＮＡＯは連絡先のリストを告げる。
【０１７４】
メールを直接送信する場合、またはそうでなければ受信メッセージに応答する場合、ロボットはユーザのメッセージを記録することになる。
【０１７５】
メッセージを終了すると、ロボットは記録されたメッセージを再読し、次に例えば以下の様々な相互作用を提案する選択ボックスを開始する。
－それを再生する／メッセージを再生する／私のメッセージを再生する：ＮＡＯはメッセージを再読する。
－メッセージを再記録する／私のメッセージを再記録する／それを再記録する。メッセージは、最初のものが適切でなければ再記録することができる。
－それを送信しない／送信しない／メッセージを送信しない。ＮＡＯはメッセージを送らず、アプリケーションの前のレベルに戻ることになる。
－それを送る／メッセージを送る／私のメッセージを送る。ＮＡＯはメッセージを送る。
－タイムアウトの場合、メッセージが送られる。
－「出る」要求または繰り返しに伴う失敗などのタイムアウトでないボックスから「他の」出る場合、アプリケーションは前のレベルに戻る。
【０１７６】
パラメータは主メニューの選択ボックスのものとほぼ同じであり、パラメータ「無返答時の最大繰り返し回数」は１に設定される。何回の無返答の後にロボットはタイムアウトがあったと考えるかを示すパラメータ「音声認識タイムアウト」と、「確認時の音声認識タイムアウト」は、ユーザが何も言わずにメッセージを容易に送ることができるように例えばデフォルト設定の６秒の代わりに４秒に設定することができる。
【０１７７】
選択ボックスはまた、ボックスの使用の継続時間全体にわたって一定であるパラメータにより静的なやり方で構成することができる。しかしながら質問を自動的に生成するシステムの利用の枠組み内では、パラメータは自動的に調整されてもよい。例えば、Ａｓ　Ａｎ　Ａｎｇｅｌ社により開発されたものなどの会話型エージェントの利用の枠組み内では、前記エージェントは、それが自動的に生成することになる、質問－回答の機能として選択ボックスを構成することができる。
【０１７８】
特に本発明の実施を容易にするように他の改良が、Ｃｈｏｒｅｇｒａｐｈｅビヘイビア開発ソフトウェアに対しなされた。以下にその説明を行う。
【０１７９】
Ｃｈｏｒｅｇｒａｐｈｅボックスは、支援されるプログラミング言語の１つのスクリプトによって実施される。このボックスが、繰り返し回数、ロボットにより利用される言語、ロボットが話さなければならないテキストなどのいくつかのパラメタータ化可能な態様を有する場合、これらの情報の項目はボックスのスクリプト中に直接組み込まれる。ボックスのパラメータの修正が望まれる場合、例えばそれを違ったやり方で使用するためにそれを複製した後、その行動を変えるためにはボックスのスクリプトを修正する必要がある。これは、使用されるスクリプト言語の完全な知識の無いユーザが実行することを望むありふれた操作であって、Ｃｈｏｒｅｇｒａｐｈｅユーザの生産性を改良するためのもので

あるので、ボックススクリプトを構成することができるように特別のインターフェースが開発された。この機能には２つの態様がある。
【０１８０】
　Ｃｈｏｒｅｇｒａｐｈｅインターフェースでは、ユーザは、ボックスの入力と出力を生成することができるのと同じやり方で、ボックスの属性を編集するためのウィンドウ内に「ボックスパラメータ」を生成する可能性を有する。各「ボックスパラメータ」は、名前、説明、タイプ（ブーリアン、整数、浮動小数点、文字列の中から）を有する。各「ボックスパラメータ」はタイプの関数としてデフォルト値などの追加属性を有することができる。最後に、「ボックスパラメータ」は親ボックスから継承するものとして定義されてもよく、これは、値が決定されるやり方に影響を与えることになる。「ボックスパラメータ」が定義されると、ボックスは、その左下角の追加可視指標によりそのチャート内に表示される。ユーザがこのアイコンをクリックすると、「ボックスパラメータ」編集ダイアログが開き、ユーザは「ボックスパラメータ」の属性内に定義される随意的制約条件の枠組み内で、各「ボックスパラメータ」に関連する値を定義することができる。
【０１８１】
　ボックスのスクリプトでは、ボックスの著者は、今後、引き数として「ボックスパラメータ」と称するいくつかの機能の助けを借りて「ボックスパラメータ」にアクセスすることができる。ボックスの著者は「ボックスパラメータ」の現在値を調べて、それを変更することができる。また、ボックスの著者は、Ｃｈｏｒｅｇｒａｐｈｅ内に現われないがボックスのスクリプト内の一時記憶として機能することができる動的「ボックスパラメータ」を生成することができる。パラメータの現在値は、親ボックスから継承したものか否かマーキングされることに依存する。それでない場合（デフォルトの場合）、「ボックスパラメータ」はボックスに固有である。ボックスのスクリプトがそれを調べると、その現在値が単純に返される。パラメータの現在値が、継承するとしてマーキングされると、値の読み込み中に、同じ名前の「ボックスパラメータ」を含む親ボックスが見つかるまでボックスチャートの階層中の逆追跡（ｂａｃｋｔｒａｃｋｉｎｇ　ｕｐ）が発生する。何も見つからなければ現在のボックスの現在値が使用される。
【０１８２】
　さらに、ロボットがそのカメラの視界に入る対象物を認識できるようにするソフトウェアモジュールが、ロボットに利用可能である。但し、認識される対象物は、最初に学習段階において学ばれなければならない。この学習はＣｈｏｒｅｇｒａｐｈｅ内の特定インターフェースの助けを借りて実行される。
【０１８３】
　このインターフェースは、ロボットのカメラにより送られる映像を実時間で表示する。画像は、Ｃｈｏｒｅｇｒａｐｈｅがカメラと正しく構成された映像キャプチャーモジュールとを有するロボットに接続された時だけ、利用可能である。映像ディスプレイが活性化されると、ユーザは学習をトリガすることができる。次に、カウントダウンが画像上に現われる。このときユーザはカメラの前の対象物を提示するのに例えば４秒を有する。カウントダウンの終わりに、画像が捕捉され記録される。次に、ユーザは、固定画像上にポリゴンを描くことにより画像内の興味のある対象物の輪郭を描かなければならない。ポリゴンが閉じられると、ダイアログが開き、対象物を定義するキーワードを入力するようユーザに依頼する。
【０１８４】
　各学習は、Ｃｈｏｒｅｇｒａｐｈｅによりユーザコンピュータ上に保存されるデータベース内にエントリを生成する。学習が終了すると、ボタンはロボットにデータベースの縮小版を送ることができるようにする。このとき、物体認識モジュールはこのデータベースを使用することになる。対象物が認識されると、関連キーワードを含む事象がロボット上でトリガされる。
【０１８５】
　Ｃｈｏｒｅｇｒａｐｈｅはまた、ロボットの行動のエディタである。図４に対する注釈

として先に説明したように、行動は、ロボットにより実行され得るコンピュータプログラムと同様なオブジェクトである。これらの行動をインストールしロボット上で実行するために、ロボット上の行動を管理するためのインターフェースの開発が行われた。Ｃｈｏｒｅｇｒａｐｈｅがロボットに接続されると、アプリケーションのメニューのエントリにより行動マネージャを表示できるようにする。これは、これらを操作するための一組のボタンだけでなくロボット上にインストールされる行動のリストも表示するモーダルウィンドウである。

【０１８６】

インストールされた行動毎に、その名前と、その状態（現在実行中か実行中でないか）と、ロボットが始動されると行動が実行されなければならないかどうかを定義する属性と、が表示される。行動を開始または停止するためには、その現在状態を表示するアイコンをクリックすることで十分であり、その効果は状態をトグル切り替えすることである。行動が終了されると、状態は自動的に元に切り替えられ「停止される」。属性「スタートアップ開始」はチェックマークが付けられるボックスである。これは属性の現在値を示し、ユーザはこの値を変えるためにこれを単純にクリックすることができる。

【０１８７】

行動のリストと共に表示されるボタンは、行動のいくつかを追加すること、行動のいくつかを削除すること、行動のいくつかをユーザコンピュータに転送することができるようにする。したがってユーザは、ロボット上にインストールされた行動を、あたかもそれらが彼のコンピュータ上のファイルかのように、極めて容易に操作することができる。特に、ユーザは、行動を、ロボット上に記録する必要無く、ダウンロードし、修正し、彼のコンピュータ上に再インストールすることができる。

【０１８８】

このときユーザによりインストールされた行動は、時間的一致の制約条件下で、および様々な行動ボックス、行動フレーム、およびタイムラインにより定義された行動間で並行して実行することができる。

【０１８９】

上に述べた例は本発明の実施形態の例示として与えられた。これは決して本発明の分野を限定するものではなく、本発明の分野は以下の特許請求範囲により規定される。

【図４】

図４

【図５】

図５

【図６】

図６

【図７ａ】

図７a

【図７ｂ】

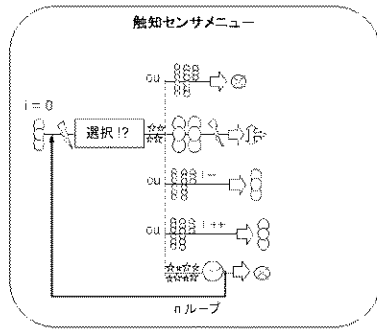



図 7b

【図７ｃ】

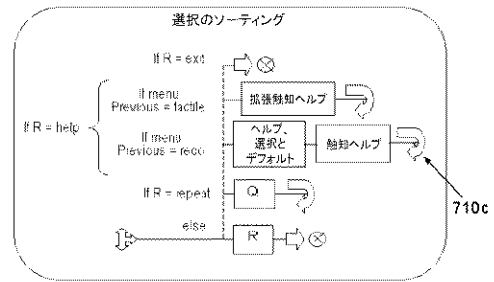



図 7c

【図８ａ】

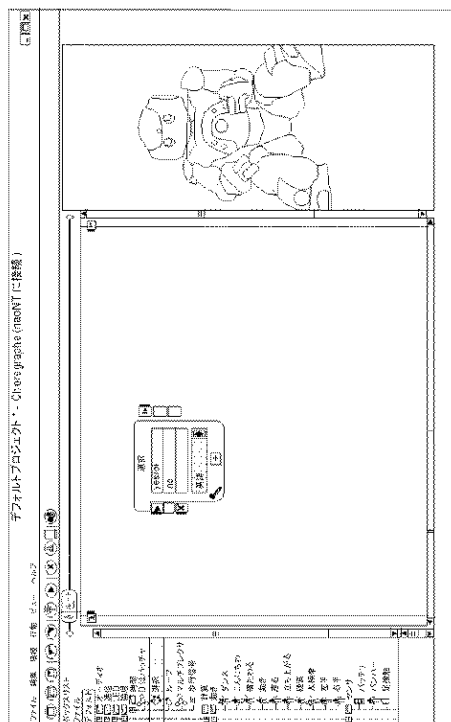



図 8a

【図８ｂ】

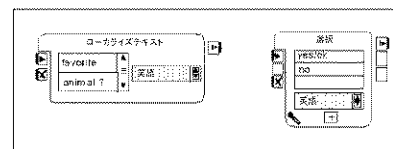



図 8b

【図８ｃ】

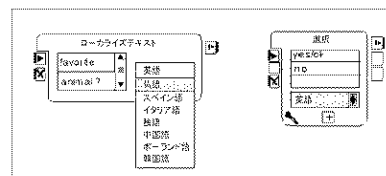



図 8c

【図８ｄ】

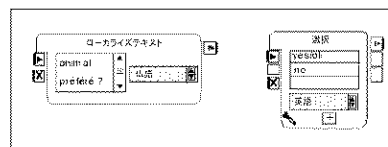



図 8d

【図８ｅ】

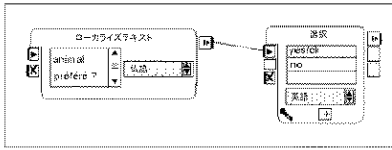



図 8e

【図９ａ】

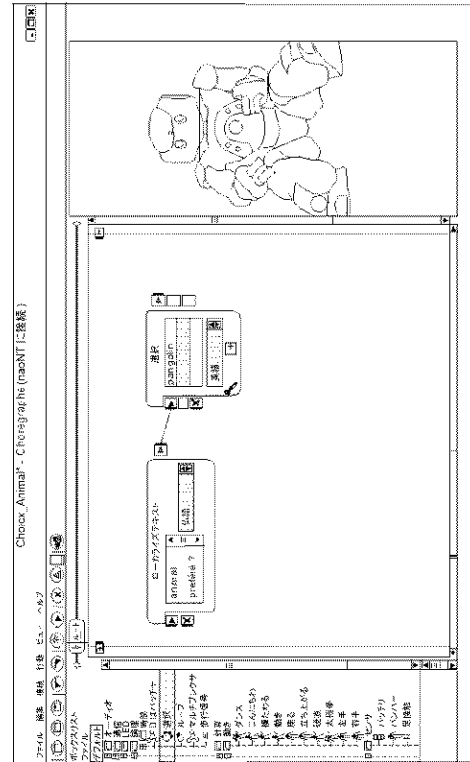



図 9a

【図９ｂ】

図 9b

【図９ｃ】

図 9c

【図９ｄ】

図 9d

【図９ｅ】

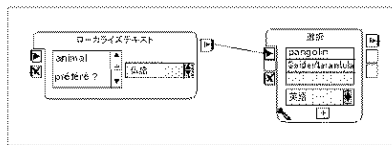
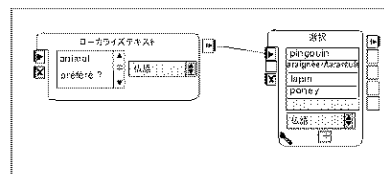
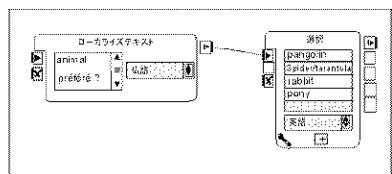
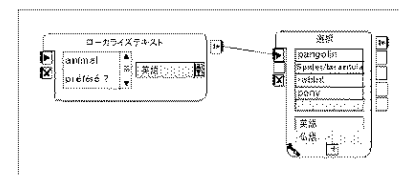
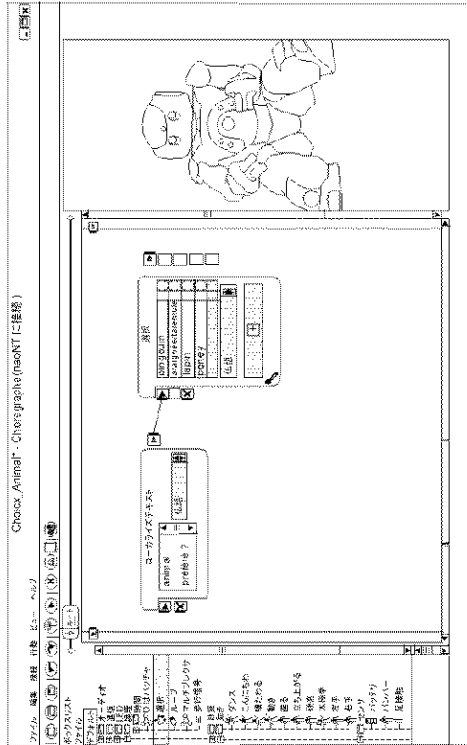
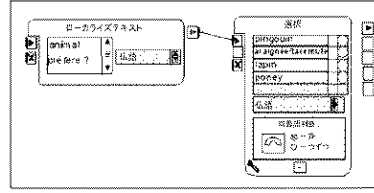
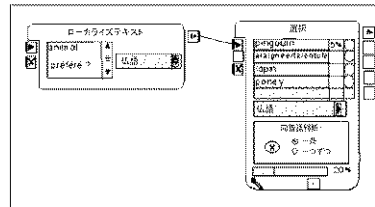
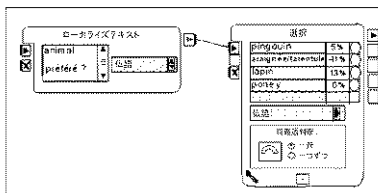
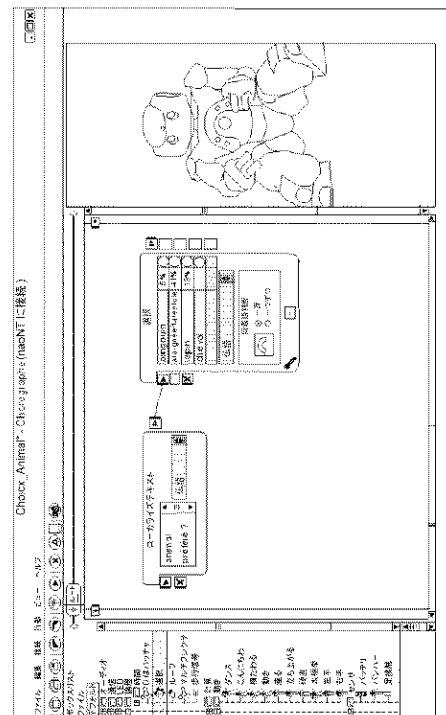
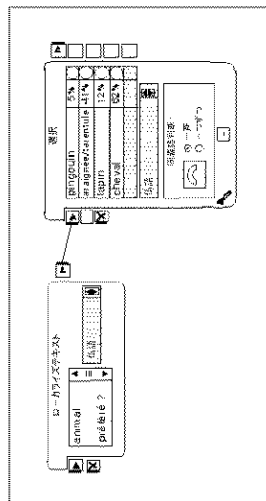
図 9e

【図１０ａ】

【図１０ｂ】

【図１０ｃ】

【図１０ｄ】

【図１１ａ】

【図１１ｂ】

図 11b

【図１２ａ】

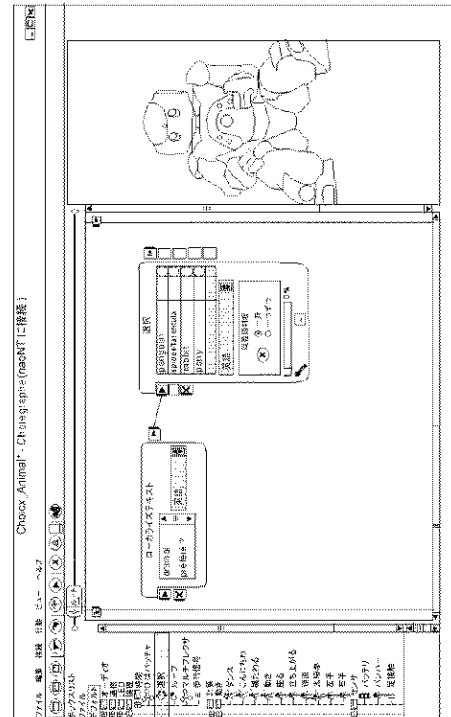



図 12a

【図１２ｂ】

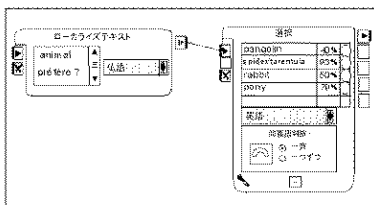



図 12b

【図１２ｃ】

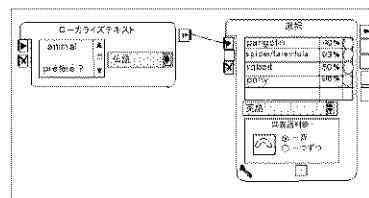



図 12c

【図１２ｄ】

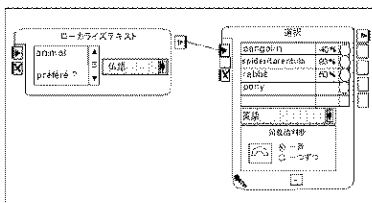
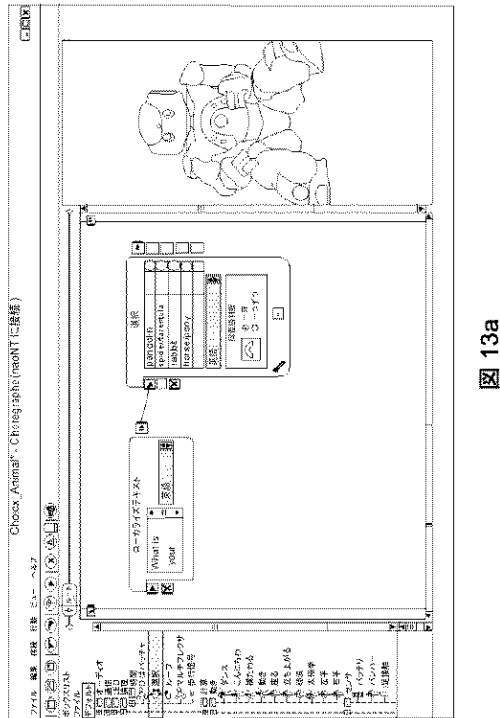
図 12d

【図１３ａ】

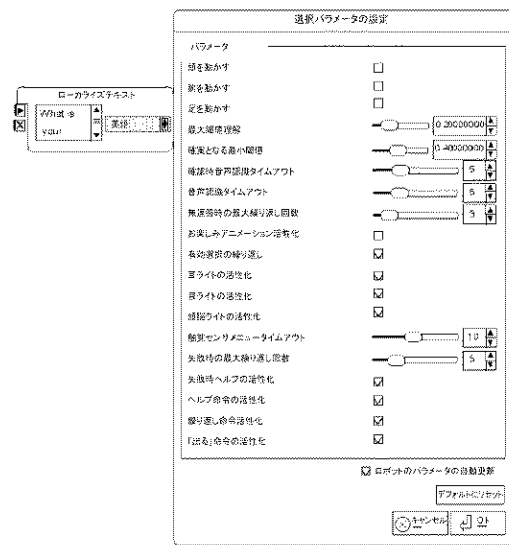
図 13a

【図１３ｂ】

図 13b

【図１３ｃ】

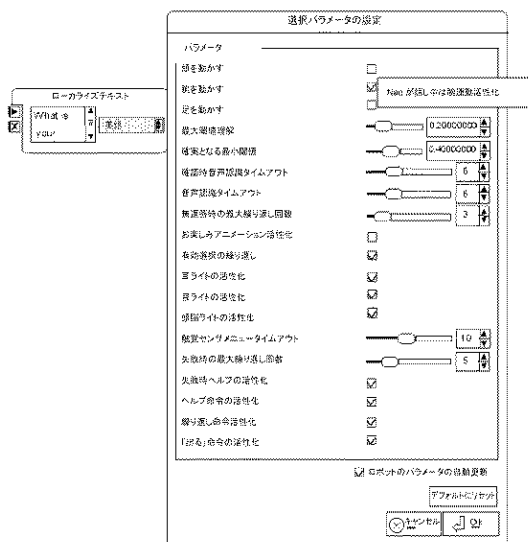



図 13c

【図１３ｄ】

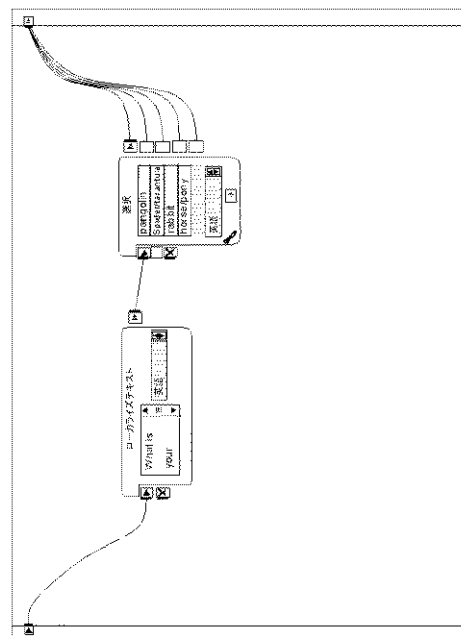



図 13d

【図１】

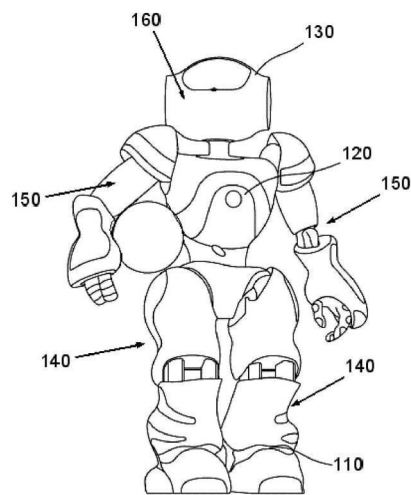



FIG.1

【図２ａ】

FIG.2a

【図２ｂ】

FIG.2b

【図３】

FIG.3